1. Report No. NASA CR-2149	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle DEVELOPMENT OF A STANDARDIZED BATTERY OF PERFORMANCE TESTS FOR THE ASSESSMENT OF NOISE STRESS EFFECTS		5. Report Date January 1973
		6. Performing Organization Code
7. Author(s) George C. Theologus, George R. Wheaton, Angelo Mirabella, Rae E. Brahlek, and Edwin A. Fleishman		8. Performing Organization Report No.
9. Performing Organization Name and Address American Institutes for Research, Washington Office, 8555 Sixteenth Street Silver Spring, Maryland 20910		10. Work Unit No.
		11. Contract or Grant No. NAS 1-9854
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D. C. 20546		13. Type of Report and Period Covered Contractor Report
		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

The primary objective of this project was to develop, construct, and test a prototype standardized human test battery for use in assessing the effects of stress on human performance. As a result of an extensive review of the factor-analytic research on human performance which has been reported in the scientific literature, a set of 36 relatively independent categories of human performance were identified. These categories encompass human performance in the cognitive, perceptual, and psychomotor areas, and include diagnostic measures and sensitive performance metrics. Then a prototype standardized test battery was constructed, and research was conducted to obtain information on the sensitivity of the tests to stress, the sensitivity of selected categories of performance degradation, the time course of stress effects on each of the selected tests, and the learning curves associated with each test. A research project utilizing a three factor partially repeated analysis of covariance design was conducted in which 60 male subjects were exposed to variations in noise level and quality during performance testing. Effects of randomly intermittent noise on performance tests showed consistent stability. The results of 14 analyses of covariance of the data taken from the performance of the 60 subjects on the prototype standardized test battery provided information which will enable the final development and test of a standardized test battery and the associated development of differential sensitivity metrics and diagnostic classificatory system. Fu-ther research will be necessary in order to reduce the size of the standardized test battery, establish reliability estimates, measure the effects of repeated trials, and develop the final testing format.

17. Key Words (Suggested by Author(s)) psychomotor test battery human performance standardized test battery assessment of noise effects on man	18. Distribution Statement Unclassified - Unlimited

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 124	22. Price* $3.00

*For sale by the National Technical Information Service, Springfield, Virginia 22151

TABLE OF CONTENTS

DEVELOPMENT OF A STANDARDIZED BATTERY OF PERFORMANCE TESTS FOR THE ASSESSMENT OF NOISE STRESS EFFECTS

By George C. Theologus, George R. Wheaton, Angelo Mirabella
Rae E. Brahlek, and Edwin A. Fleishman
American Institutes for Research

SUMMARY

The primary objective of the first year of this project was to develop, construct, and test a prototype standardized test battery for use in assessing the effects of stress on human performance. As a result of an extensive review of the factor-analytic research on human performance which has been reported in the scientific literature, a set of 36 relatively independent categories of human performance were identified. These categories encompass human performance in the cognitive, perceptual, and psychomotor areas, and include diagnostic measures and sensitive performance metrics. Then a prototype standardized test battery was constructed, and research was conducted to obtain information on the sensitivity of the tests to stress, the sensitivity of selected categories of performance degradation, the time course of stress effects on each of the selected tests, and the learning curves associated with each test. A research project utilizing a three factor partially repeated analysis of covariance design was conducted in which 60 male subjects were exposed to variations in noise level and quality during performance testing. Effects of randomly intermittent noise on performance of the reaction time tests were observed, but most of the other performance tests showed consistent stability. The results of 14 analyses of covariance of the data taken from the performance of the 60 subjects on the prototype standardized test battery provided information which will enable the final development and test of a standardized test battery and the associated development of differential sensitivity metrics and diagnostic classificatory system. Further research will be necessary in order to reduce the size of the standardized test battery, establish reliability estimates, measure the effects of repeated trials, and develop the final testing format.

I. INTRODUCTION

Complex man-machine systems often subject man to a wide range of unusual environments and to an even greater variety of environmental stressors. Consequently, it is becoming increasingly important to understand how these environments and stressors affect human performance and, ultimately, the performance of the system itself. Because of the significance of this problem, there has been a great deal of research on the effects of such unusual environments as isolation and confinement, high altitude, prolonged space flight, and undersea habitation. Accompanying stressors such as hypoxia,

sleep loss, noise, and vibration also have been studied. In spite of these efforts, however, there has been little generalization of results beyond the specific problems studied, and little payoff for either applied or basic research.

Examination of the literature on environmental stressors and their effects upon performance reveals a number of problems impeding the development of a systematic body of knowledge and the application of that knowledge. Within any particular study, environmental stressor variables and performance measures may be defined precisely, and employed effectively. However, it is often impossible to relate them systematically to different variables from other studies. Various researchers have employed different measures of performance, different methods of testing, and different data analysis techniques. Descriptive data relating to the values and ranges of environmental stressors employed, the intervals of exposure, and the characteristics of subjects indicate considerable variation from study to study. Until the relationships among these and similar components are established systematically, development of a cohesive body of knowledge may be impossible.

The experiences of Tune (1964) in attempting to synthesize and interrelate the results of studies on the effects of high altitude are typical of the difficulties which one encounters. He found the results of numerous studies to be either contradictory or to be reported in insufficient detail to permit conclusions. The results varied for different devices and research methods to the point that Tune concluded: "It is impossible to compare the performance of _Ss_ in these studies because of lack of standardization of the techniques used" (p. 560). Repeated statements to this same effect make it clear that development and use of a standardized battery of performance tests, together with appropriate experimental procedures, is essential to the growth of knowledge about environmental stressors, and to the effective application of that knowledge. As one step toward developing the required body of knowledge, many investigators are now stressing this approach. They are becoming aware of the tremendous advantages of programmatic and integrative research which transcends the highly specific and often isolated studies which have characterized behavioral research in the past.

They emphasize at least three advantages of paramount importance. First, programmatic research with standardized procedures would permit systematic comparison of the effects of environmental stressors from one study to another. Second, it would permit the effects of environmental stressors to be mapped for different and distinct aspects of human performance. Third, this integrative approach eventually might permit the prediction of stressor effects from laboratory studies to "real-world" tasks.

To cope with the problem of developing a rigorous method for systematically assessing performance, a number of strategies have been adopted. While strikingly different on the surface, they share the common objective of providing for an incisive evaluation of performance. Strategies at one end of the continuum have involved development of sophisticated, high-fidelity simulators which attempt to capture each nuance of operational settings. While simulators and other complex equipments are undoubtedly necessary for developing and evaluating operator proficiency in many man-machine systems, it may be more difficult to justify their use for the assessment of operator performance as a function of selected stressor variables. When concern focuses on the effects of various stressors (confinement, noise, vibration, ambient temperature, etc.), special testing facilities are required. One cannot reasonably expect to provide simulations of such diverse projects as SST, APOLLO, LEM, and MOL and to equip each with the special facilities and equipment required to study a broad range of stressors. Moreover, even if such specialized facilities were available, one is still faced with the problem of systematically comparing the results of studies conducted within these unique and diverse systems.

In light of these difficulties, several prominent researchers (Alluisi, Fleishman, French, Guilford, and Parker) have suggested an alternative approach. In this approach the attempt is made to distill complex performance into more fundamental units of behavior. By focusing on parameters of behavior assumed to underlie performance in all man-machine systems, their evaluation procedure has the potential for permitting the comparison of research results across systems and among studies. This approach is predicated upon the development of a standardized battery of performance tests and the methodology associated with its use.

The present report describes the initial efforts in a long-range program of research designed to develop a standardized performance battery (SPB) for assessment of the effects of environmental stressors on human performance. In the next section of this report, Section II, the general objectives and philosophy of the overall research program are discussed. Section III describes the research activities which were carried out in the effort to develop the prototype SPB. Topics discussed include the isolation and definition of the basic parameters of human performance which served as a foundation for the SPB, and the selection of specific tests of these performances. Also discussed are reviews of the human performance and noise effects literatures, together with a diagnostic methodology for guiding the assessment, interpretation, and organization of research results on the effects of environmental stressors. The fourth section of the report contains a detailed description of a research study which was conducted in order to provide preliminary data on the sensitivity of selected tests in the SPB to noise effects. The fifth and final section of the report summarizes the research effort and presents the general conclusions stemming from this project.

II. BACKGROUND

The long-range objective of this program of research is to construct a standarized battery of performance tests through which statistical predictions and subjective inferences can be made from knowledge of the effects of stressors on these tests to the effects of those stressors upon "real-world" tasks. The research program will generate a test battery through which a stressor's "behavioral profile" may be obtained. The test battery will comprehensively measure the basic parameters of human performance in a stress-sensitive fashion.

This objective is founded upon the rationale that different components of human task performance will be affected differentially as a function of the type and level of stressor under which the performance is produced, and that measurement of these changes will provide a quantitative evaluation of the disruptive effects of the stressor and of man's vulnerability to it.

Approaches to Performance Assessment

In the investigation of environmental stressors, the concept of performance assessment must be broad-based, permitting consideration of a wide range of possible behavioral changes. It is recognized that while a particular stressor may be studied with respect to one or two modes of performance, it may often modify many other aspects as well. For example, the unanticipated and often unwanted side effects of drugs make it necessary to consider many other aspects of behavior in addition to the particular response of interest. In many instances, this broad consideration is a prerequisite to complete understanding of the effects of the agent or stressor employed. General conclusions about the behavioral effects of a stressor certainly cannot be made in terms of a single index. When many aspects of performance are of interest, as they must be for a generally meaningful program of evaluation, a wide range of behaviors must be measured.

Concern for broad assessment arises because there is no single or universal measure of performance. Performance does not represent a single substantive entity. Therefore, of the many modes of behavior which might be considered, some manageable subset has to be selected. Two main strategies have been employed in selecting measures of performance. First, some investigators have chosen to utilize a single apparatus permitting single or multiple measures of relatively complex performance. Second, other investigators have suggested the use of performance test batteries. Within this approach, one group of investigators has employed synthetic task situations hypothesized to reflect rationally-derived critical "functions" typically performed by operators in man-machine systems. An additional group of investigators has emphasized an empirically-based performance battery approach to assessment.

Single Apparatus Approach

A rather common practice in many laboratories is the systematic use of a single testing apparatus to obtain measures of performance under varied stressor conditions. Such devices, comprising complex tracking tasks, for example, are ordinarily quite sophisticated, permitting experimental variation of display control elements, feedback loops, and related dynamics. In light of the variations which are possible, performance on these tasks can best be described as complex. That is, the performance measure or measures which are monitored arise from the complex interaction and integration of many specific and diverse aspects of behavior.

Two major problems frequently accompany use of the single apparatus approach as opposed to the more broadly-based approaches. First, there is the problem of selecting the device around which to organize an assessment program. Devices often are sought which, in some global sense, are "sensitive" to the range of stressors being studied. Therefore, measures which do not show decrement as a function of selected stressors are often rejected as "insensitive." This practice is not compatible with comprehensive research programs in which the objective

is to determine the full range of effects. In such comprehensive programs it is just as important to know the particular kinds of tasks which are unaffected as it is to identify tasks on which performance is degraded or enhanced.

The second problem concerns the diagnostic power associated with the single apparatus approach. Research on a diversity of stressors has repeatedly revealed differential effects for different aspects of performance. For example, Gorham and Orr (1957, 1958) investigated the effects of experimentally-induced stress (fear) on the components of a complex skill. The general finding of this research was that different skill components (measured by a test battery) were affected differentially by particular treatments. For example, performance on perceptual-vigilance and numerical ability components deteriorated, while motor co-ordination and response orientation components were maintained, despite the same exposure to stressors. Were one to have worked with a single complex apparatus, such differential effects would not have been easily deduced. For example, if the apparatus were based upon a tracking task, knowledge that total integrated tracking error increased under stress would provide little diagnostic information. It would be difficult to determine whether degradation on the total task stemmed from degradation on the perceptual component, the motor component, or both. Although of use in other testing contexts, "multi-purpose" testing equipments are hard pressed to provide the mapping of differential effects which is basic to any comprehensive program of stress evaluation.

Performance Battery Approach to Assessment

An alternative to the single apparatus approach is the identification of a relatively limited number of behavioral components which underlie performance in a wide variety of complex tasks. Tasks reflecting each of these aspects of performance could be combined into a standard performance battery. Use of the standard battery in a comprehensive program of research would: (a) permit comparisons among studies in which the battery was employed; (b) enable the mapping of differential effects

to be conducted for different aspects of performance under a wide range of stressors; and (c) eventually provide data from which predictions to a wide range of "real-world tasks" might be attempted.

The initial problem in the development of a comprehensive battery involves the decision as to how to distinguish among different aspects of performance. Not too long ago a favorite distinction was between mental and motor tasks or between cognitive and non-cognitive tasks. Such broad distinctions, however, were not very helpful in pinpointing various aspects of behavior or in developing specific testing instruments. More recently, two alternative conceptualizations have emerged from which test batteries might be developed. The first, based on task analysis, conceives of different aspects of performance in terms of certain basic and recurrent "functions" performed by operators in complex man-machine systems. The second, based on factor-analytic research, focuses on relatively independent categories of "abilities," "traits," or "capacities" which contribute to performance on a wide variety of tasks.

<u>Rationally-derived "functions" approach</u>. Categories which conceive of man-task interactions in terms of classes of "functions" provide one basis for a performance battery approach to assessment. Over the years, the rational-descriptive procedure of task analysis has given rise to a number of categories. Working with categories originally developed at Lockheed, Alluisi (1967) has abstracted these functions and embodied them in synthetic task situations. Further, he has succeeded in combining these separate tasks into a multiple-task performance test battery (MTP).

Since its development, the MTP has been fruitfully employed in a number of assessment programs. For example, it has been used to obtain measures of group performance during confinement (Alluisi, Chiles, Hall, and Hawkes, 1963) and in studying the effects of sleep loss on crew performance (Alluisi, Chiles, and Hall, 1964). More recently, the MTP has been employed in an extensive program assessing the behavioral

effects of infectious diseases (Alluisi, Thurmond, and Fulkerson, 1966). The battery was valuable for these purposes because of the differential performance assessment which it permitted.

The multiple-task performance bettery represents a major step forward in the development of human performance testing methodologies. Testing with this battery would be preferable to the single apparatus approach in the study of stressors. Tasks comprising the battery are reported to provide reliable measurements of performance. Similarly, the battery provides enough realism to maintain the interest and co-operation of subjects. Again, however, there are problems associated with its use.

The first problem concerns the choice of "functions" which are used to describe performance. The most comprehensive information about correlations among human performance measures indicates a greater degree of specificity, and considerably more diversity, than is represented by the "functions" which traditionally are chosen. Just as complex human performance cannot be represented adequately by a single testing device, it probably requires many more than a few component "functions" for its comprehensive description and, therefore, for the thorough assessment of stressor effects.

The second difficulty with the "functions" approach involves the method used to select "functions" and tests of "functions." These "functions" are derived rationally from task-analytic data. In many cases they are labels for rather general types of undifferentiated human activity. As such, their relationship to more basic and underlying parameters of performance is unknown. Consequently, how can one distinguish between activities which represent a given "function" and other activities which do not? By the same token, how can one be sure of what "function" a particular test measures? Until these questions are resolved, there can be no direct evidence that tests of these "functions" are independent. Such independence is essential to creation of a parsimonious battery of tests.

Empirically-derived "abilities" approach. The final alternative makes use of categories of human performance which have been derived from empirical studies of the correlations among human performance data. Much of what is known about the categorization of performance comes from such experimental-correlational research. These categories allow one to be much more specific about performance than do the more general categorical terms derived from the "functions" approach. Furthermore, because of an empirical factor-analytic basis, empirical indices (i.e., factor loadings) can be employed to assess the diagnosticity of a given test with reference to any particular performance category.

The experimental-correlational research which has generated these performance categories attempts to empirically sift and categorize the many tests available in the different areas of human performance. The basic approach of this research is to examine large numbers of tests from the various areas of human performance to determine what they have in common. The attempt then is made to identify the number of independent kinds of performance which are measured by the tests. Based on this approach, extensive work has been conducted in the cognitive, perceptual, physical proficiency, and psychomotor areas of performance. For example, Fleishman (1964) in his work on psychomotor tests found several hundred tests in use. Administration of these tests to large samples of subjects permitted factor analyses of the relationships among these measures. Through a series of interlocking studies, Fleishman found that actually only ten or eleven measurably distinct factors or, in his terms, "basic abilities" were being assessed by all of these tests. He was further able to identify one or two "best" tests for each of these factors so that both efficient and broad assessment of the area would be feasible.

These extensive factor-analytic research programs have provided a means for selecting specific, generally non-redundant tests within each of four broad categories of human performance: cognitive, perceptual, psychomotor, and physical proficiency. They also have indicated the

minimum number of factors which appear to function independently within each category.

Research on selected stressors already has been conducted with these prototype "ability" batteries. Evans and Consolazio (1967), using tasks from Fleishman's prior analysis of physical proficiency, studied the effects of high altitude on three kinds of physical proficiency: explosive strength, dynamic strength, and stamina. Using standard tasks (medicine ball puts, chin-ups, and a bicycle ergometer) they found that while dynamic strength was unaffected, the other two were. Explosive strength returned to normal within one week, whereas stamina did not completely recover during the course of the experiment. Prior physical conditioning was beneficial in reducing the amount of performance decrement observed.

Also, Baker, Elkin, Van Cott, and Fleishman (1966) have investigated the effects of various drugs on human performance utilizing a test battery based upon factor-analytically-derived performance categories. The battery of tests measured performance in the psychomotor, physical proficiency, cognitive, and sensory-perceptual areas. The results of the research enabled the construction of profiles of differential performance decrement. Although the test battery which was employed was far from complete, it is precisely this type of research which is required for environmental stressors. A battery of "basic ability" tests is uniquely suited to comprehensive and diagnostic evaluation of the effects of environmental stressors.

III. DEVELOPMENT OF THE STANDARDIZED PERFORMANCE BATTERY

The immediate objectives of the first year of the project were the development, construction, and testing of a prototype SPB based upon empirically-derived "human ability" categories. Further, within the first year of research, the preliminary testing of this battery was to be accomplished in a study designed to determine the effects of selected noise stressors on three tests chosen from those in the battery. The activities carried out in pursuit of these objectives will be elaborated upon in the following paragraphs.

Reviews of the Human Performance Literature

The first project activity was to review the human performance literature in the following two areas: (1) factor-analytic studies of human performance, and (2) previous research on the development of performance batteries to assess stress effects. The review of the factor-analytic literature was used as the primary vehicle for defining categories of human performance and for selecting an initial set of performance tests. The review of previous attempts to develop stress test batteries was undertaken to augment the factor-analytic review, and to provide guidance in test selection.

Within the literature dealing with the factor analysis of human performance, several hundred studies were examined. However, more than half proved to be of no value to the project and, consequently, were eliminated from further consideration. Many of these studies treated limited and esoteric experimental questions or focused on areas of performance which were not germane to the project. Others were isolated non-programmatic efforts in which no attempts were made to provide for replication or confirmation of the findings. Finally, some of the studies which were rejected contained methodological shortcomings which impaired the value of the research results. The major methodological failures present in many factor-analytic studies have been summarized by McNemar (1951) and Guilford (1952). Among these are: (1) extraction of

too many factors, (2) insufficient sample size, (3) use of spurious correlations, (4) use of tests of low reliabilities, (5) lack of experimental independence among variables, and (6) definition and interpretation of factors on the basis of minimal factor loadings.

The studies selected for more intensive review constituted a nucleus of well-executed, substantiated research in which meaningful parameters of human performance had been identified. In order to integrate the findings from these studies, the reported performance parameters (i.e., factors) were identified and extracted, and a file representing these factors was developed. Each factor was represented in the file by its name, an accompanying definition of the factor, and references to studies in which that particular factor had been identified. The factors in the file were then classified into major areas of human performance. The vast majority fell within the psychomotor, cognitive, and perceptual domains of performance. Additional factors fell within the physical proficiency and sensory areas. At this point in the project, however, the latter two sets of factors were set aside from the mainstream of analysis. The physical proficiency factors appeared to be only minimally involved in tasks of probable interest to NASA, while the sensory factors were too few in number and too ambiguously defined to warrant inclusion in the initial prototype battery.

The factors within each of the three major areas of performance were carefully examined to group those which appeared to represent the same basic dimensions of human performance. Since no convention exists for naming factors, and since factors are renamed as additional insights into their composition are acquired, the factor names themselves were of little value in the process of classification. For example, the factor of Associational Fluency has been variously referred to as Associational Fluency, Associational Speed, Factor A, and Facility with Verbal Relations. Therefore, the factor definitions and the nature of the tests loading on the factors were used as the basis for comparison and classification.

Examination of the resultant groups of factors revealed some instances in which a "group" consisted of only one factor or two remotely-related factors. In these cases there was no convincing or prepotent evidence for the existence of a stable factor. Rather, in most cases these factors seemed artifactual, arising primarily as a function of the specific test battery and type of measurement employed in the study in which they were identified. The consequent elimination of these factors produced three final sets of factors which provided a solid basis for test selection. These sets of factors are listed in Table 1, and are presented with relevant references in Appendix A.

TABLE 1

Performance Dimensions Identified in the Literature Reviews

Psychomotor	Cognitive
Control Precision	Associational Fluency
Multilimb Coordination	Expressional Fluency
Response Orientation	Ideational Fluency
Reaction Time	Word Fluency
Speed of Arm Movement	Induction
Rate Control	Associative Memory
Manual Dexterity	Mechanical Knowledge
Finger Dexterity	Memory Span
Arm-Hand Steadiness	Number Facility
Wrist-Finger Speed	Originality
Aiming	General Reasoning
	Semantic Redefinition
Perceptual	Syllogistic Reasoning
Flexibility of Closure	Sensitivity to Problems
Speed of Closure	Verbal Comprehension
Length Estimation	Figural Adaptive Flexibility
Perceptual Speed	Semantic Spontaneous Flexibility
Spatial Orientation	
Spatial Scanning	
Visualization	

Within the psychomotor area of performance the relevant factors are defined almost exclusively by the work of Fleishman and his associates (summarized by Fleishman, 1964). In the cognitive and perceptual domains the relevant factors are based upon the work of many authors, although Guilford's research (summarized by Guilford, 1967) is probably the most comprehensive in these areas.

Some confirmation of the performance dimensions which we succeeded in identifying in the cognitive and perceptual domains was obtained by comparing these dimensions with those identified by French, Ekstrom, and Price (1963). An examination of both lists, in terms of factorial definitions and the tests representing the various factors, suggested that the two lists were similar. To reduce confusion and to provide continuity with this previous research, the factor names which French and his associates employed were adopted for purposes of the present research. The similarity of the two lists was also an indication of the fact that few new contributions to this portion of the factor-analytic literature have been made since the survey by French et al. (1963).

The reviews of the literature on previous attempts at developing test batteries for the assessment of stress effects did not lead to the identification of any additional parameters of performance. In general, these previous efforts differed philosophically from the current effort. These studies attempted to represent either a job sample or a set of human "functions" or "processes" in the test batteries. We, on the other hand, attempted to isolate basic and independent parameters of human performance and to include tests representative of these parameters in the battery. For our purposes, the tests in these other batteries would have to be considered factorially impure and, consequently, incorporating them into the current battery would add an unknown degree of redundancy.

Criteria for Test Selection

Prior to the selection of tests for the performance dimensions identified in the literature reviews, criteria for test selection were

developed. Based upon a consideration of the objectives of the project, a set of 11 criteria were generated. Included in the following list are several criteria adapted from Alluisi and Fulkerson (1964).

1. Generality - The test battery must encompass as many critical aspects of human performance as possible, while simultaneously minimizing redundancy.

2. Validity - Each test selected must actually measure that aspect of human performance which it is intended to measure; each test must exhibit sufficient face validity to permit subjective generalization of the effects of a given stressor on that test to the effects of that stressor on a "real-world" task; taken as a whole, the battery must permit statistical predictions from the effects of a stressor on the laboratory tests to its effects on "real-world" tasks.

3. Reliability - Each test must exhibit high reliability (e.g., test-retest and internal consistency).

4. Standardization - Each test must be compatible with standardized administration under controlled conditions.

5. Sensitivity - Each test must be capable of reflecting the genuine performance impairments which may occur under the effects of a stressor.

6. Engineering feasibility - Equipment suitable for programming, presenting, and partially scoring performance must be available at reasonable cost; it must be reliable in its operation.

7. Flexibility - Each test must be modifiable to permit reprogramming for a range of stimuli.

8. Trainability - Asymptotic levels of performance should be quickly attainable during pre-stress administration.

9. Safety - The test equipment should not present a potential health or safety hazard to volunteer subjects.

10. Maintainability - Test equipment must be readily maintainable and not vulnerable to damage by stressed volunteers.

11. Ease - Each test should be easily set up, administered, and scored.

Selection of Performance Tests

Based upon the reviews of the performance literature, tests meeting the selection criteria were chosen to represent each of the primary performance dimensions which had been identified. The process of test selection required comparisons among the various tests which had relatively high loadings on the relevant factors. In the psychomotor area, the process of test selection was facilitated by Fleishman's specification of appropriate tests for each of the factors which he identified (Fleishman, 1964). This process was further facilitated by the availability of the Perceptual-Motor Performance Console (PEMCON). This device was developed under contract to NASA (NASA 2542) by Biotechnology, Inc. Development of the device was founded on Fleishman's research in the psychomotor performance area, and was guided by a consideration of the kinds of activities performed by astronauts as derived from a task analysis of the Gemini Mission. A detailed description of PEMCON and of the tests comprising it is presented in a report to NASA by Biotechnology, Inc. (1966). PEMCON provides 18 tests of perceptual-motor skills. Included in these tests are measures of ten of the eleven psychomotor abilities shown in Table 1. A measure of the Aiming factor is not included. Fleishman (1964) specifies the Pursuit Aiming test for this factor, and an additional three tests for this factor have been identified by Fleishman and Hempel (1954). These latter tests are Square Marking, Small Tapping, and Marking Accuracy.

In the cognitive and perceptual areas, the selection of tests was aided by the synthesis performed by French et al. (1963). This synthesis was invaluable since few researchers other than Guilford have programmatically attempted to develop and refine measures for the cognitive and perceptual factors which they have identified. Without such synthesis, any attempt on our part at test selection would have involved a detailed and laborious examination of each of the tests which loaded highly on

each of the factors identified in the various studies. We also would have had to evaluate each test with respect to criteria such as factorial purity, availability of alternate forms, redundancy, level of difficulty, ease of administration, and reliability. Therefore, to avoid the possible expenditure of unnecessary time and effort in cognitive and perceptual test selection, the sets of tests developed by French et al. (1963) were examined. The review of these tests and the related selection criteria demonstrated that further test selection efforts by AIR would have been redundant. For this reason, the set of tests identified by French et al. were utilized.

In their original selection of tests, these authors employed panels of experts in human performance assessment who evaluated the available tests against several criteria. "The tests for each factor were selected so as to be : (a) three in number, (b) such as to provide for covering as much of the range from sixth grade through college as possible, (c) as factorially pure as possible for the intended factor, (d) as different as possible to balance out uniqueness, and (e) reasonably easy to administer by paper-and-pencil methods" (French et al., 1963, pp. 2-3).

The only major criterion which these authors did not employ was test reliability. To acquire this information both for these tests and for the psychomotor tests, a review was undertaken of studies in which these tests were originally employed and of subsequent studies which utilized these tests for research purposes. This review indicated that little reliability information was available and that the information which was available was primarily based upon split-half estimates. Table 2 presents the reliability data which were available. These data indicate that the tests on which data have been collected are quite reliable. However, the data are rather incomplete. Split-half and alternate-forms reliability estimates are unavailable for many of the tests, and test-retest estimates are absent for all of the tests.

TABLE 2

Reported Reliabilities for Tests in the Prototype Standardized Performance Battery

Test	r	Reference
Factor: Aiming		
Square Marking	.92*	Fleishman & Hempel (1954)
Small Tapping	.89*	Fleishman & Hempel (1954)
Marking Accuracy	.91*	Fleishman & Hempel (1954)
Factor: Associational Fluency		
Associational Fluency I	.44*	Guilford & Hoepfner (1966)
Associations IV	.62**	Guilford & Christensen (1956)
Factor: Expressional Fluency		
Expressional Fluency	.66*	Guilford & Hoepfner (1966)
Simile Interpretation	.70*	Guilford & Hoepfner (1966)
Word Arrangements	.70**	Guilford & Christensen (1956)
Factor: Induction		
Letter Sets	.64*	Lemke, Klausmeir & Harris (1967)
Locations	.82*	Lemke, Klausmeir & Harris (1967)
Factor: Associative Memory		
Picture - Number	.76*	Duncanson (1966)
Object - Number	.79*	Duncanson (1966)
First and Last Names	.81*	Duncanson (1966)
Factor: Memory Span		
Auditory Number Span	.58*	Duncanson (1966)
Letter Span - Auditory	.68*	Duncanson (1966)
Factor: Number Facility		
Addition	.91*	Duncanson (1966)
Division	.92*	Duncanson (1966)
Subtraction & Multiplication	.90*	Duncanson (1966)
Factor: Originality		
Plot Titles (clear)	.74**	Guilford & Christensen (1956)
Symbol Production	.86**	Guilford, Berger & Christensen (1955)
Consequences (remote)	.74**	Guilford, Kettner & Christensen (1956)
Factor: Perceptual Speed		
Finding A's	.81*	Duncanson (1966)
Number Comparison	.79*	Duncanson (1966)
Identical Pictures	.88*	Duncanson (1966)

*Split-half estimate
**Alternate forms estimate

TABLE 2 (Cont.)

Test	r	Reference
Factor: General Reasoning		
Mathematic Aptitude	.44*	Duncanson (1966)
Ship Destination	.93*	Lemke, Klausmeir & Harris (1967)
Necessary Arithmetic Operations	.74*	Lemke, Klausmeir & Harris (1967)
Factor: Semantic Redefinition		
Gestalt Transformation	.51*	Guilford, Wilson & Christensen (1952)
Object Synthesis	.72*	Guilford, Wilson & Christensen (1952)
Picture Gestalt	.41*	Guilford, Wilson & Christensen (1952)
Factor: Syllogistic Reasoning		
Nonsense Syllogisms	.88***	Lemke, Klausmeir & Harris (1967)
Logical Reasoning	.72*	Lemke, Klausmeir & Harris (1967)
Factor: Sensitivity to Problems		
Apparatus	.89**	Guilford, Kettner & Christensen (1955)
Seeing Problems	.78*	Guilford & Hoepfner (1966)
Seeing Deficiencies	.58**	Guilford, Berger & Christensen (1955)
Factor: Vocabulary		
Vocabulary (V-1)	.62*	Duncanson (1966)
Vocabulary (V-2)	.68*	Duncanson (1966)
Factor: Figural Adaptive Flexibility		
Match Problems II	.70*	Guilford & Hoepfner (1966)
Match Problems V	.64*	Guilford & Hoepfner (1966)
Factor: Semantic Spontaneous Flexibility		
Utility	.42*	Guilford & Hoepfner (1966)
Alternate Use	.62*	Guilford & Hoepfner (1966)
Object Naming	.47**	Guilford, Frick, Christensen & Merrifield (1957)

*Split-half estimate
**Alternate forms estimate
***Kuder-Richardson 20 estimate

One additional test was selected for inclusion in the prototype SPB. An examination of the performance dimensions listed in Table 1 revealed that time-sharing performance was not included. The primary reason for its absence is that factor-analytic studies have not examined this aspect of performance. The review of the stress literature indicated that divided-attention tasks were particularly sensitive to stress effects and, consequently, were important in any attempt to adequately assess stress. Further, Alluisi (1964) included an attentive function in a specification of the major performance requirements of man-machine systems. On these bases, the dimension of Attention was added to the list of performance parameters in Table 1. Tests of this dimension are available in PEMCON.

The result of these test selection efforts was the development of an initial form of the SPB. The candidate tests for inclusion in the prototype SPB are shown in Table 3. As can be seen in this table, several tests are available for performance assessment within each category of performance. Additional research in later phases of this project will have to be devoted to the selection of the "best" tests for the categories of performance. Selection of tests for each category will be made so as to maximize the independence of the final set of tests.

Review of the Noise Literature

The review of the noise literature was undertaken in order to determine how noise affects performance and to suggest procedures by which these effects could best be measured on the tests selected for inclusion in the SPB. During the course of this review, it became necessary to also review selected areas of the general stress literature. As a result of this secondary review, an approach was formulated for organizing the literature on noise effects in a diagnostic fashion. The eventual format for this organization will be based on an integration of the performance degradation taxonomy presented by Chambers (1963, p. 301) with the categories of performance incorporated in the terminal SPB. It will provide a two-dimensional framework within which to organize

TABLE 3

Candidate Tests Included in the Prototype
Standardized Performance Battery

Factor	Tests
Psychomotor	
Control Precision	Rotary Pursuit** Two-dimensional, one-hand pursuit tracking (PEMCON)*
Multilimb Coordination	Rudder Control** Two-Hand Coordination** Complex Coordination** Two-dimensional, two-hand compensatory tracking (PEMCON)*
Response Orientation	Discrimination Reaction Time** Direction Control** Choice Reaction Time** Four-position choice reaction time task (PEMCON)*
Reaction Time	Reaction Time** Visual and auditory reaction time tests using key press response (PEMCON)*
Speed of Arm Movement	Two Plate Tapping** Moving arm horizontally between two switches (PEMCON)*
Rate Control	Single Dimension Pursuit** Rate Control** Motor Judgment** Two-dimensional, one-hand compensatory tracking (PEMCON)*
Manual Dexterity	Minnesota Rate of Manipulation** Modification of Minnesota Rate of Manipulation using a single object (PEMCON)*

*Parker, Reilly, Dillon, Andrews, and Fleishman (1965)
**Fleishman (1964)

TABLE 3 (Cont.)

Factor	Tests
Finger Dexterity	O'Connor Finger Dexterity** Purdue Peg Board** Modification of standard tests, using a minimum of components (PEMCON)*
Arm-Hand Steadiness	Track Tracing** Holding stylus in hole without touching sides (PEMCON)*
Wrist-Finger Speed	Tapping** Rapid alternate tapping of two keys using one hand (PEMCON)*
Aiming	Pursuit Aiming** Square Making*** Small Tapping*** Marking Accuracy***
<u>Perceptual</u>	
Flexibility of Closure	Hidden Figures**** Hidden Patterns**** Copying****
Speed of Closure	Gestalt Completion**** Concealed Words****
Length Estimation	Estimation of Length**** Shortest Road**** Nearer Point****
Perceptual Speed	Finding A's**** Number Comparison**** Identical Pictures**** Comparison of two meter indications (PEMCON)*
Spatial Orientation	Card Rotations**** Cube Comparisons**** Spatial Orientation****

*Parker, Reilly, Dillon, Andrews, and Fleishman (1965)
**Fleishman (1964)
***Fleishman and Hempel (1954)
****French et al. (1963)

TABLE 3 (Cont.)

Factor	Tests
Spatial Scanning	Maze Tracing**** Choosing a Path**** Map Planning****
Visualization	Form Board**** Paper Folding**** Surface Development****
Cognitive	
Associational Fluency	Controlled Associations**** Associational Fluency I**** Associations IV****
Expressional Fluency	Expressional Fluency**** Simile Interpretations**** Word Arrangements****
Ideational Fluency	Topics**** Theme**** Thing Categories****
Word Fluency	Word Endings**** Word Beginnings**** Word Beginnings and Endings****
Induction	Letter Sets**** Locations**** Figure Classification****
Associative Memory	Picture-Number**** Object-Number**** First and Last Names****
Mechanical Knowledge	Tool Knowledge**** Mechanical Information**** Electrical Information****
Memory Span	Auditory Number Span**** Digit Span**** Letter Span****

****French et al. (1963)

TABLE 3 (Cont.)

Factor	Tests
Number Facility	Addition**** Division**** Subtraction and Multiplication****
Originality	Plot Titles**** Symbol Production**** Consequences****
General Reasoning	Mathematics Aptitude (1)**** Mathematics Aptitude (2)**** Ship Destination**** Necessary Arithmetic Operations****
Semantic Redefinition	Gestalt Transformation**** Object Synthesis**** Picture Gestalt****
Syllogistic Reasoning	Nonsense Syllogisms**** Logical Reasoning**** Inference****
Sensitivity to Problems	Apparatus**** Seeing Problems**** Seeing Deficiencies****
Verbal Comprehension	Vocabulary (1)**** Vocabulary (2)**** Extended Range Vocabulary**** Advanced Vocabulary (1)**** Advanced Vocabulary (2)****
Figural Adaptive Flexibility	Match Problems II**** Match Problems V**** Planning Air Maneuvers****
Semantic Spontaneous Flexibility	Utility**** Alternate Uses**** Object Naming****

****French et al. (1963)

information about stress effects. For purposes of the present study, the noise literature was evaluated only with respect to the performance degradation dimension. Classification on only one dimension was necessary because of uncertainty as to which performance categories may comprise the terminal SPB.

To aid in this process, Chambers' taxonomy was reorganized into five areas, each of which incorporates several degradation subcategories:

1. Changes in performance distribution over time (lapses," "queuing," and "escape").

2. Changes in performance distribution among tasks, among task components and within tasks ("falling off," omissions," "filtering," and "fixation-block-confusion").

3. Increases in performance rate ("performance oscillation," "inadvertent control inputs," and "sudden changes in rate or frequency").

4. Decreases in performance rate ("changes in phasing" and "response lags").

5. Changes in performance accuracy ("increases in error amplitude," "failure to detect," "error in returning, integrating, steering, and processing information," "disassociation," "stereotyping," "approximations," and "perceived disintegration.")

Our efforts in organizing the noise literature in terms of the above framework are summarized in tabular form in Appendix B. A major finding of this review and categorization is that noise stressor research has been limited to seven of the 19 subcategories of performance decrement. That is, only one-third of the potential forms of performance degradation have been examined vis-a-vis noise stress. Further, the bulk of the research has been concerned with Categories 4 and 5, i.e., decreases in performance rate and increases in error.

The review of the noise stressor research also indicated that noise stressors may have their most potent effects on performance

metrics other than mean performance level. For example, based upon the relatively small amount of research relevant to Categories 1 and 2 in Chambers' taxonomy (Appendix B), there is some indication that noise effects might be most pronounced on performance distributions across time and across tasks. Taken as a whole, the variety of effects which have been found within the different degradation categories indicate that effective diagnosis of the effects of noise and other environmental stressors is not predicated solely upon the possession of a set of tests which adequately covers the range of human performances. Diagnosis of stress effects also requires a set of performance measures comprehensive and sensitive enough to reveal the subtleties of human performance. Many researchers are becoming aware that insensitive analytic and measurement techniques are responsible for much of the inconsistency and ambiguity in research on the effects of noise stress. Man's capability to compensate for stress effects through mechanisms such as queuing or redistribution of performance across time, tasks, or task components may render gross measures such as mean performance level inadequate for the diagnosis of stress effects. This finding is one of the more salient encountered during the course of this research. In light of it, we have included in the long-range project goals the development and assessment of performance metrics for a variety of categories in the performance degradation taxonomy.

IV. TEST OF THE STANDARDIZED PERFORMANCE BATTERY

Following the development of a prototype form of the SPB, an experimental study was conducted of the effects of two noise stressors on three tests from the battery. This research was designed to provide preliminary information regarding (1) the sensitivity of tests in the battery to noise effects, (2) the sensitivity of performance metrics for selected categories in Chambers' performance degradation taxonomy, (3) the time course of noise effects on each of the selected tests, and (4) the learning curves associated with each of these tests.

Methods and Procedures

Subjects. The subjects employed in this study were 60 male students obtained from area universities. They were paid volunteers, recruited with the aid of advertisements in their school newspapers. The average subject was 21.18 years of age (S.D. = 2.95). All subjects were screened for hearing defects by means of self report.

Experimental tasks. Three experimental tasks chosen from the tasks contained in PEMCON were utilized in this study. The PEMCON system consists of an operator's console (Appendix D) and a separate experimenter's console. A detailed description of this device can be found in a report to NASA by Biotechnology, Inc. (1966). Briefly, the operator's console contains a cathode ray tube (CRT), a pair of control sticks for one- or two-hand tracking, and a variety of meters and back-lighted push buttons. The experimenter's console contains an array of push buttons and rotary switches for programming experimental tasks. It also includes a pair of meters, a clock, and a counter for recording subject responses. For the present research this equipment was placed in two rooms, separated by a one-way mirror. The experimenter and his console were located in one room and the subject and his console were in the other.

The first experimental task was the Reaction Time task. In this task the subjects were required to keep the index finger of their non-preferred hand on a back-lighted push button and to press the button as

rapidly as possible whenever the light behind the button came on. The rate of stimulus presentation was eight per minute, with stimulus onsets spaced at approximately equal intervals. This task reflected the Reaction Time factor and yielded measures of response latency recorded on a Hunter electronic clock (Model 120A).

The second task was the Rate Control task. This task required the subjects to track, in a compensatory fashion, a point of light which moved away from the center of a CRT, under sine wave, first-order dynamic programming. To track the dot of light, the subjects manipulated a control stick mounted on a universal joint. Movement of the control stick simultaneously affected the x and y axes of the CRT through a pair of independent potentiometers. The subjects were instructed to track the point of light with their preferred hand. This task represented the Rate Control factor and provided measures of integrated tracking error recorded on galvanometers for the x and y axes.

Lastly, a Time Sharing (divided-attention) task was employed to measure the factor of Attention. This task required the subjects to perform the above two tasks simultaneously. That is, the subjects had to perform the Reaction Time task with their non-preferred hand while simultaneously tracking the point of light with their preferred hand. The combination of the Reaction Time and Rate Control tasks was chosen for the Time Sharing task because it was felt that this combination would provide a difficult but not impossible divided-attention task. Preliminary data collected on eight subjects supported this belief. The use of these two task components in the Time Sharing task also permitted examination of the effects of noise on the Reaction Time and Rate Control factors both in isolation and in combination.

Independent variables. The manipulated variables were noise quantity and noise quality. The levels of noise quantity were 0 dB (re: 0.0002 dynes/cm^2) and 85 dB. The noise quality levels were randomly intermittent noise and patterned intermittent noise. The randomly intermittent noise consisted of 85 dB of broad-band white noise

delivered in bursts of random duration, varying in nine one-second steps between one and nine seconds. The interpolated intervals of silence were likewise random, varying in three one-second steps between one and three seconds. This noise averaged five seconds on and two seconds off. The patterned intermittent noise consisted of 85 dB of broad-band white noise presented in five-second bursts with interpolated two-second quiet intervals.

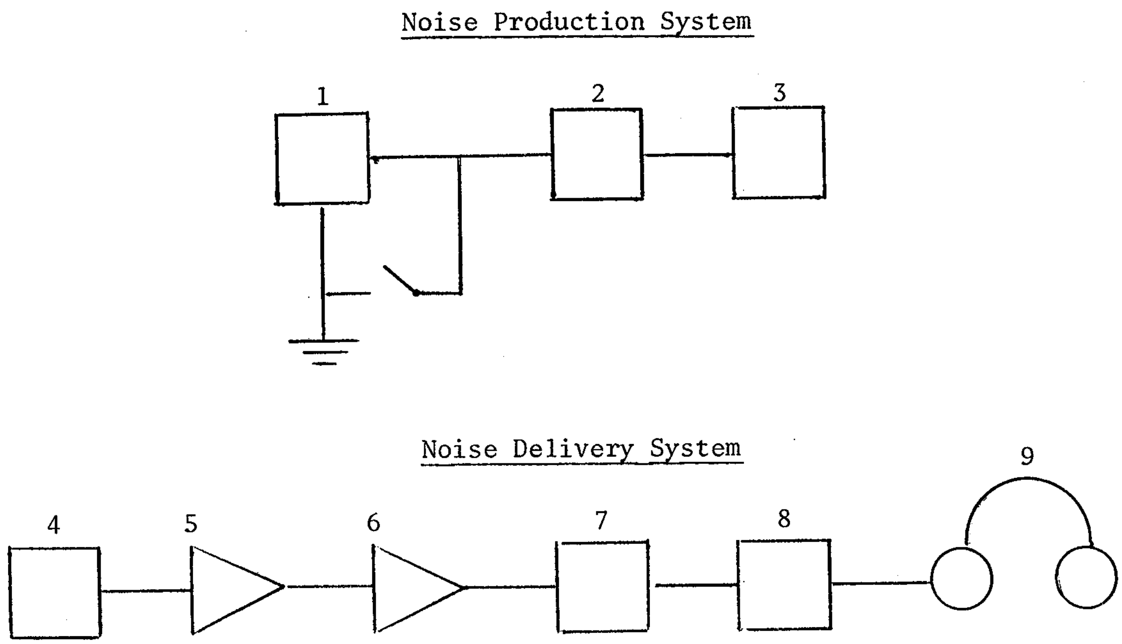
These noise stressors were delivered to the subjects through a set of Grason-Stadler earphones (Model TDH-39) from a Viking tape deck (Model 230). Intensity at the earphones was controlled by a Daven step attenuator which contained its own plug-in matching transformers. The entire system, outlined in Figure 1, was calibrated with a B & K earphone calibrator.[1] The resulting frequency response curves are shown in Table 4. This table includes the open filter response which was used to define the reference intensity at the attenuator.

The noise tapes were originally developed from a Lafayette noise generator (Model 15012) and a Bogen power amplifier (Model AP 250). The noise intermittency was programmed with a microswitch which shorted the output of the noise generator. This circuitry is also block-diagrammed in Figure 1.

The choice of the specific noise stressor conditions was guided by considerations of safety and validity. Since a relatively prolonged exposure to noise was planned for this research, care had to be taken to insure the safety and well-being of the experimental subjects. A guideline to safe noise exposure is provided by the data of Kryter, Ward, Miller, and Eldredge (1966). As can be seen in Table 4, the 85 dB intensity level employed in this study is well within the safe limits at all frequency levels.

---

[1]The calibration of the noise delivery system and development of the noise tapes were carried out at the Forest Glenn Annex of Walter Reed Army Medical Center in the Audiology and Speech Center with the assistance of Mr. M. B. Whitlock and Mr. L. Teague.



1. Lafayette Model 15012 White Noise Generator
2. Bogen Model AP 250 Power Amplifier
3. Tape
4. Viking Model 230 Tape Deck
5. Viking Model RP 83 Pre-amplifier
6. Viking Model PA 94B Power Amplifier
7. Daven Step Attenuator
8. 500 ohm Matching transformers from attenuator
9. Grason-Stadler Model TDH-39 earphones

Figure 1. Diagram of noise generation system.

TABLE 4

Frequency Response of Sound Generation System
Measured at Each TDH-39 Earphone

Band (Hz.)	Center (Hz.)	Safe Level (dB)*	Measured Level (dB)**	
			Right phone	Left phone
22.1-44.2	31.5		59	49
44.2-88.4	63.0		62	59
88.4-177.0	125.0	118	68	66
177.0-354.0	250.0	107	71	68
354.0-707.0	500.0	99	69	70.5
707.0-1414.0	1000.0	95	75	74.0
1414.0-2828.0	2000.0	91	77	77.5
2828.0-5657.0	4000.0	90	87	87.5
5657.0-11314.0	8000.0	95	78.5	78.5
11314.0-22628.0	16000.0		61.5	56.0

*Kryter et al. (1966); based upon a single one-hour exposure per day for 10 years.

**Based upon open filter reading of 90 dB for the right phone and 89 dB for the left phone.

The second consideration in the choice of stressors was validity. A major and recurrent problem in stress research is the lack of a universally-accepted definition of stress. Researchers typically have ignored this problem and manipulated conditions of ad hoc interest. They generally have chosen extreme stress levels; for example, very high G forces, noise in excess of 100 dB, and temperatures which deviate widely from accepted comfort levels. This approach, however, was not reasonable for the current effort, since the thrust of the effort was upon test development and not upon evaluation of environmental effects per se. Assurance was needed that the environmental conditions which were chosen were in fact "stressful." Otherwise, it would be difficult to interpret negative results. That is, without an independently defined stressor, we could not be certain as to whether an

absence of noise effects was due to non-stressful conditions or to insensitivity of the experimental tasks to a stressful noise condition.

The solution to this problem was to define the stressor in terms of its physiological and psychological effects. The physiological effects of moderate intensity white noise have been documented by Oken, Heath, Shipman, Goldstein, Grinker, and Fisch (1966) who have shown that it has the potential for increasing EMG, heart rate, skin conductance, and systolic blood pressure. The psychological effects of moderate intensity noise have been demonstrated by Atherly, Gibbons, and Powell (1970), and Sullivan, Warren, and Dabice (1970). Further, unpredictable noise has been shown to produce performance decrement and to overload human information channels (Sanders, 1961; Finkelman and Glass, 1969; McCann, 1969; and Wyon, 1970).

Experimental design. The design chosen for this research was a partially repeated measures design with repeated measures over blocks of trials and the noise quality factor. This design is shown in Figure 2. In this design the 60 experimental subjects were randomly assigned to one of six groups (n = 10). Two groups of subjects performed on each of the experimental tasks. Within each task one group was assigned to the 0 dB noise quantity level and the other to the 85 dB condition. This design permitted efficient use of the experimental subjects and provided for a reduction of inter-subject variability within noise quantity levels. High inter-subject variability in reaction to stress normally is expected. Greater control over inter-subject variability could have been achieved with a completely repeated measures design. Such a design, however, would have been severely confounded with order effects. These effects would have involved both the order of presentation of the tasks and of the noise stressors. In the partially repeated measures design which was chosen, order effects due to presentation of tasks were eliminated, since the repeated measurement was not across tasks. A fixed order of presentation of the noise stressors was utilized.

Task	Subject Group	Noise	Noise Quality	
			Blocks of Trials	Blocks of Trials
Reaction Time	1	Quiet	Quiet	Quiet
	2	85 dB	Random	Patterned
Rate Control	3	Quiet	Quiet	Quiet
	4	85 dB	Random	Patterned
Time Sharing	5	Quiet	Quiet	Quiet
	6	85 dB	Random	Patterned

Figure 2. Experimental design.

Daily testing procedure. On each day of testing, four subjects reported to AIR's Washington Office at half hour intervals beginning at 8:30 A.M. Upon arrival the subjects filled out personal data sheets and then waited in a lounge until called to the experimental room for the first of three sessions of testing. Each session lasted approximately one half hour. There was a 1-1/2 hour break between sessions. Thus, the subject's experimental day lasted approximately 4-1/2 hours. Following completion of the last session, each subject was paid in cash for his participation and dismissed.

The purpose of the first testing session was to familiarize each subject with the task to which he had been assigned. This step was taken to reduce variance due to learning in data collected in the later experimental sessions. This session also provided covariate data on each subject which were used in the data analyses to adjust scores for initial between group differences in performance.

During the first session each subject was brought individually to his work station, seated at the console, and given a set of instructions relevant to his particular task. The operation of the operator's console then was demonstrated. Once the subject indicated understanding of his task, the familiarization trials were begun. During these trials the subject was instructed to put on the earphones although no noise stimulus was presented. In this session, 20 trials were administered, each trial lasting one minute. These trials were spaced rather than massed with one-minute rest pauses given at the end of each block of five trails. It was felt that this procedure would help stabilize baseline performance more rapidly. For the Reaction Time task, eight presentations of the stimulus light were administered within each one-minute trial. In the Rate Control task, tracking performance was continuous over the one-minute trial. The Time Sharing task combined both of these conditions within each one-minute trial. From the subject's point of view performance within and between trials was uninterrupted. In actuality, the experimenter had to read and reset the clock between

each reaction time presentation, and he had to read and reset the integrated error meters between each trial.

The second and third testing sessions were utilized to gather the experimental data. Prior to the second session each subject was assigned either to the experimental condition (85 dB) or to the control condition (0 dB). During the second session, each experimental subject was exposed to the randomly intermittent noise stressor while he was performing his particular task. During the third session, patterned noise was presented. The control subjects received no noise stimulation in either the second or third sessions although they were required to wear the earphones. The testing procedures for the second and third sessions were identical to those for the first session with the following two exceptions. First, two one-minute practice trials were administered prior to the 20 trials in each experimental session. Second, the 20 experimental trials were massed. That is, the one-minute breaks, which were given after each block of five trials in the familiarization session, were not given in the two experimental sessions.

Data Analysis

As a result of conduct of this experiment the following basic data were obtained. For each subject assigned to the Reaction Time and Time Sharing tasks, 160 response latencies, measured in milliseconds, were recorded for each experimental session. For each subject who performed on the Rate Control and Time Sharing tasks, 20 integrated error scores were recorded in each experimental session for the x and y tracking axes. These basic data were not subjected to analysis. Rather, they provided a basis for derivation of dependent measures specifically tailored to provide information on the issues of primary concern: differential sensitivity of selected degradation categories, the time course of noise effects, and the amount of learning over experimental sessions.

Data reduction. In order to provide information regarding the time course of noise effects and the extent of learning across sessions, stable estimates of subjects' performance were required for different time segments. Accordingly, the decision was made to collapse the data within each session into four blocks of five trials. This procedure yielded eight data points per subject. Inspection of the raw data indicated that this blocking procedure served to reduce rather high trial to trial variability noted in the data of some subjects. All subsequent dependent measures were derived from these blocked data.

Derivation of dependent measures. One of the primary reasons for execution of the research was to determine whether selected categories from the degradation taxonomy would exhibit differential sensitivity to noise stress. Consideration of the basic data suggested a wide variety of dependent measures which could be used to represent each of the five major categories in the taxonomy. Faced with a large number of potentially relevant measures it was therefore decided to limit initial concern to those rather commonly used to describe performance on reaction time and tracking tasks. During the anticipated second phase of the research, additional and more esoteric measures will be examined. Those included in the present analysis represented the following three degradation categories:

1. Changes in performance distribution over time--assessed by frequency of response blocks and mean duration of response blocks;

4. Decreases in performance rate--assessed by mean response latency and reaction time; and

5. Changes in performance accuracy--assessed by mean integrated tracking error.

The dependent measures used to evaluate Categories 1 and 4 were derived from the response latency data obtained in the Reaction Time and Time Sharing tasks. Category 5 was assessed by measures based on the integrated error scores from the Rate Control and Time Sharing tasks.

The first of the two measures used to represent Category 4 was the mean response latencies for each block of trials. However, examination of individual subjects' response latencies revealed a problem which precluded direct analysis of these data. At random points in time within subjects' records, responses were present which were substantially longer than what was apparently a "normal" latency. These longer latencies had two effects. First, they produced a severe skewing of the distributions, and second, some of these longer latencies caused subjects to miss the immediately following trial or trials.

In order to normalize the distributions of response latencies, a reciprocal transformation of the data was employed. Subsequent inspection of the transformed distributions revealed that the skewness had been substially reduced. For each subject the mean of the reciprocal response latencies was calculated at each block of trials. The missing data points caused by the longer than normal response latencies did not present a major problem. Since the data were to be treated in blocks of trials, the mean for any block in which a data point was missing was calculated on the reduced number of data points recorded within that block. The means of these measures for each group are presented in Table 5. This table can be found in Appendix C, as can all tables and figures relating to this experiment.

The appearance of the "longer than normal" response latencies in the data suggested that, in addition to variation in response time, a lapse-in-performance phenomenon was being measured. It was assumed that these longer responses were related to momentary lapses in attention during which a subject failed to detect presentation of the stimulus. This lapse-in-attention phenomenon is referred to in the psychological literature as response blocking (Bills, 1931, 1935, 1937). Since no universally acceptable quantitative definition of a response block is presently available, an operational definition which had proved useful in a study of long-term visual monitoring (Baker and Theologus, 1970) was adopted. Therefore, for the present study, a

response block was defined for each subject as any latency which was two or more times longer than the mean of his 20 fastest response latencies. For each subject the definition of a response block was based upon his 20 fastest trials in the familiarization session. This session was used to define a response block since no subject received any experimental treatment during this period.

Given a definition of a response block for each subject, two measures relevant to Category 1 were derived. First, for each subject a frequency count of response blocks within each block of trials was made to develop the measure of frequency of response blocks. Second, for each subject the mean duration of all response latencies defined as response blocks was calculated within each block of trials to create the measure of mean duration of response blocks. The means of these two measures for each group are shown in Tables 6 and 7.

Examination of these three derived measures revealed that the measure of mean response latency evaluated the total distribution of scores for each subject while the measures of frequency of response block and mean duration of response blocks evaluated the scores in the upper tail of the distribution. To investigate changes in performance on scores which fell within the subject's "normal" response range, a fourth measure was developed. Since all scores two or more times longer than the mean of a subject's 20 fastest response latencies were defined as response blocks, we defined all of the scores falling below this criteria as reaction times. Therefore, for each subject the mean of all scores less than a response block was calculated within each block of trials to generate the measure of mean reaction time. This constituted the second measure used to evaluate Category 4. The group means for this measure are presented in Table 8. The use of this measure in combination with the two measures relating to response blocking permitted us to determine whether any noise effects manifest in the mean response latency measure were due to an increase in the number or duration of "longer than normal" scores or to a change in the subject's "normal" speed of response.

In order to examine Category 5 from the performance degradation taxonomy, the integrated error scores obtained on the Rate Control and Time Sharing tasks were employed. For each subject, three measures were obtained within each block of trials. At each block of trials, total mean integrated error on both tracking axes, mean integrated error on the x-axis, and mean integrated error on the y-axis were calculated. Tables 9, 10, and 11 present the group means for these measures.

Analytical procedures. The primary means for the analysis of the various performance measures was analysis of covariance. These analyse were performed on an IBM 1130 computer system, using a program specifically developed for these data. For each of the performance measures the covariate employed to adjust the measures from the two experimental sessions was the similar measure calculated for each subject on the last ten trials from the familiarization session. For example, the covariate for frequency of response blocks was the frequency count for each subject of the response blocks made during the last 80 presentations of the reaction time stimulus in the familiarization session. The mean covariates for each group on each measure are presented in Tables 5 through 11. Since a single covariate for each subject was used to adjust all eight of his experimental data points, only the between subjects terms in the analysis of covariance were adjusted. With a single covariate the within subjects regression coefficient is equal to zero, and thus, there is no adjustment of the within subjects terms. With respect to the factor and cell means, only those which reflected a between subjects factor were affected by the covariate adjustment.

Our initial intention in the calculation of the analyses of covariance was to employ a 2 x 2 x 2 x 4 (Tasks x Noise Levels x Noise Qualities x Blocks) design for each dependent measure. However, an inspection of the variances associated with each of the measures, as they were represented in Time Sharing task and in the relevant single task, indicated that variances in the Time Sharing task were far

larger. In order to avoid problems arising from severe heterogenity of variance within the analyses of covariance, the dependent measures were analyzed separately for each task. The designs which were analyzed were 2 x 2 x 4 (Noise Levels x Noise Qualities x Blocks) designs. Fourteen such analyses of covariance were carried out. These analyses are shown in Tables 12 through 25.

Following the analyses of covariance, single degree of freedom *F*-tests were calculated to investigate the issues of primary experimental interest. The effects of noise on performance for each of the dependent measures were determined through comparisons between the 0 dB and 85 dB groups at each block of trials in each session. The amount of learning which occurred between the two experimental sessions was ascertained by comparisons between the control group (0 dB) means for Session 2 and Session 3. All of the one degree of freedom *F*-tests were calculated on the adjusted means derived from the analyses of covariance. These means for each task and each measure are graphically illustrated in Figures 3 through 16.

Results and Discussion

The results of the analyses of covariance demonstrated that there were no significant main effects associated with the noise level factor on any of the dependent measures. However, the *a priori* comparisons between the experimental and control groups at each block of trials in both sessions revealed the presence of significant noise effects on the Reaction Time task but not on the other two tasks. For the measure of mean response latency, those subjects who were exposed to the random noise while performing on the Reaction Time task, produced significantly slower response latencies at the first [F (1, 143) = 7.12, p <.01] and second [F (1, 143) = 6.90, p <.01] blocks of trials. Significant random noise effects on this task also were found on the measure of mean reaction time. Similar to the effects on mean response latency, the randomly intermittent noise significantly slowed the mean reaction

times of the experimental subjects at both first [$F_{(1, 143)} = 8.06$, $p < .01$] and second [$F_{(1, 143)} = 6.36$, $p < .05$] blocks of trials. Since no random noise effects were demonstrated in the Reaction Time task for either of the two measures related to response blocking, the effects of random noise on mean response latency can be interpreted as indicating that randomly intermittent noise slows the speed at which the subjects "normally" respond but does not increase the number or duration of "longer than normal" responses (response blocks).

These results provide substantial support for the approach which we have taken to the assessment of stress effects. This approach is founded on the rationale that different types and levels of stressors will variously affect different components of human task performance, and that these effects may become manifest only in certain categories of performance degradation. The present results indicate that random, but not patterned, intermittent noise selectively affects Reaction Time performance and that this effect is detectable only through performance measures associated with Category 4 in the performance degradation taxonomy. Although further research with a wider range of stressors, measures, and tests is required to confirm these results, the obtained differential sensitivity provides an excellent indication of the potential of the SPB for effective diagnosis of stress effects and for derivation of "behavioral profiles" for various stressors. Further, the specificity of the current results with regard to both components of performance and degradation categories underscores the complexity of noise effects and suggests that previous ambiguous findings on noise effects should be reviewed in light of the more precise and diagnostic methodology presented in this report. Much of the inconsistency and ambiguity in the noise literature may be due to the use of insensitive analytic and measurement techniques.

Although no further significant effects of noise could be demonstrated in this research, a consistent but statistically non-significant degradation of performance can be noted in the Time Sharing task.

Under randomly intermittent noise, performance apparently was affected on both the primary task (Rate Control component) and the secondary task (Reaction Time component). On the primary task, mean integrated error scores for the x-axis, the y-axis, and both axes combined show and increasing degradation in performance over the four blocks of trials in Session 2 (Figures 11, 13, and 15). On the secondary task, a corresponding effect is present for the measures of mean response latency, mean reaction time, and frequency of response blocks (Figures 3, 5, and 7). No effect can be observed on the measure of mean duration of response blocks (Figure 9). Although the observed decrements in performance are not statistically significant, their consistency suggests the possible presence of a random noise effect on the performance category of Attention. Further, they suggest that the performance decrement increased over time. This is opposite to the significant effects on the Reaction Time task in which the decrement decreased over the blocks of trials in Session 2 (Figures 4 and 6).

Some partial support for these inferences can be found in two previous studies of noise stress. Finkelman and Glass (1969) investigated the effects of 80 dB periodic and aperiodic white noise on a divided attention task, comprised of a single dimension, first order tracking task and a digit recall task. The authors found that the aperiodic noise stressor significantly degraded digit recall performance. While no significant effects were found on the tracking task, the obtained differences were similar to those in the present research. The subjects who performed under aperiodic noise had higher error scores than those who performed in the presence of periodic noise. The results of the present research also resemble the results of a study by Sanders (1961). As was mentioned above, the observed degradation in performance on both components of the Time Sharing task increased over the four blocks of trials in Session 2. This observation corresponds to Sanders' demonstration that the effects of a randomly changing noise become apparent only after 15 minutes of performance under noise.

There are three possible reasons for our failure to detect a significant effect of random noise on the Time Sharing task. First, it is possible that performance was not assessed over a sufficiently long period of time. Since the decrement was observed to increase over the half hour performance period, testing beyond the half hour limit may have revealed significant noise effects at later points in time. The problem concerning the time course of stress effects has appreciable impact on the development of a final form of the SPB. To guarantee the diagnosticity of the SPB, each test in the battery must be administrable over a period of time long enough to insure that given stressors have achieved their maximum effect. Due to the importance of this issue, future research with the SPB will have to be conducted to specify a standard administration time for the tests in the battery.

A second possible reason for the failure to significantly establish an effect of random noise on the Time Sharing task is the rather high inter-subject variability which was obtained for most dependent measures on this task. This variability inflated the error terms in the analyses of covariance and the one degree of freedom *F*-tests, and thereby reduced the potential for detecting significant effects. It is not unlikely that this variability stemmed from differential responsivity of the subjects to the noise stressor. Such differences classically have been obtained with a wide variety of stressors. One means for reducing the effects of this variance on statistical tests is to employ a covariate which reflects inter-subject differences in reaction to stress. Covariates such as heart rate, blood pressure, skin conductance, and EMG are available for this purpose and should be examined in the later phases of this research program.

Finally, the lack of statistical confirmation for the apparent effect of random noise on the Time Sharing task may be due to the limited selection of performance metrics which was employed. Other measures may be more sensitive to the effects of randomly intermittent noise. Since the review of the noise literature discussed in Section III

of this report revealed that noise may have its most potent effects on the distribution of performance across time and across tasks, further analysis of the present data should be undertaken. The data from the two components of the Time Sharing task should be analyzed for changes in performance distribution across tasks, and the integrated error data need to be examined with respect to changes in performance distribution over time. These and other analyses will be carried out during the planned second phase of the research program.

The last issue of research interest in the present study concerned the amount of learning which occurred between experimental testing sessions. To examine this issue, one degree of freedom *F*-tests were calculated between the mean performance levels of the control groups at the second and third sessions. No significant differences were obtained on any of the measures. On this basis, it can safely be concluded that learning on the tasks was essentially complete by the end of the familiarization trials and that it did not continue over the experimental sessions.

V. SUMMARY AND CONCLUSIONS

The research activities completed during the current phase of the project have yielded three major products. The primary research product was the development of a prototype form of the SPB. This test battery is based upon a set of 36 relatively independent categories of performance which were identified through a thorough search of the literature on human performance. In light of the best information available from the factor-analytic research on human performance, these 36 categories exhaustively cover the range of performances in the cognitive, perceptual, and psychomotor areas. To assess performance in these categories, three or four candidate tests for each category have been selected for possible inclusion in a final form of the test battery.

The experience gained during the initial developmental efforts on the SPB revealed several areas which will require additional research before a final form of the battery can be derived. First, in its present form the SPB would require 36 tests for the complete assessment of performance in the categories which have been tentatively selected. In many situations the administration of so large a test battery would be impractical. Thus, to enhance the utility of the SPB, some reduction in the size of the battery, without concommitant reduction in the scope of the performances encompassed, would appear beneficial. Such a modification could be accomplished through the elimination of any tests which exhibit a degree of redundancy. An examination of the candidate tests which have been selected and of the factor analytic studies upon which they are based, reveals that some overlap apparently exists among the cognitive and perceptual tests. Therefore, these tests in particular will require some experimental examination to assure their independence. The removal of redundancy from the tests comprising the SPB would reduce the number of tests necessary for comprehensive performance assessment and would result in substantial savings in time and effort in the later use of the SPB in experimental and applied setting

Additional research is also required to establish reliabilities for the tests included in the SPB. The review of available reliability data revealed that adequate estimates have not been established for many of the tests. Two types of reliabiliy are of primary importance to the development of the SPB and will have to be established in the next phase of the research program. To insure that performances of subjects on a test are comparable and that unambiguous interpretation of scores on the various tests is possible, an estimate of the homogeneity of items within each test is required. These estimates are given by split-half reliabilities which provide information on the internal consistency of the tests under consideration. In addition to being internally consistent, the score on each test must be stable over time. Temporal stability of test scores is required in order to guarantee that the tests are providing true estimates of a subject's performance, and that they are not reflecting random daily fluctuations in the subject's condition. These latter estimates are provided by test-retest reliabilities.

Third, some future research efforts need to be devoted to modification of the form of the tests included in the SPB. At present we can foresee that some tests selected for the final form of the SPB will require modification in order to be amenable with standardized and facile administration, and with repeated presentation. The modifications primarily will entail development or elimination of test items for those tests which are shorter or longer than some standard administration time. A standard administration time is necessary to obtain comparable performance periods across tests. This becomes critical in those situations where tests from the SPB are utilized to assess acute stress effects. Under these conditions, varying performance periods across tests would make for difficult interpretation of differential stress effects. Selection of a common administration time will be guided by additional research on the time course of noise effects and by considerations pertaining to the overall time for administration of the entire test battery. In addition, to allow for repeated administration of tests in the SPB, development of new items for some tests may be

required. Many of the cognitive and perceptual tests are based upon items which are "perishable". That is, a subject can be exposed to these items only once before complete learning occurs and the item loses its value. For these tests, alternate forms would be necessary to permit repeated measurement.

Not only must the administration time for each test in the SPB be standardized, but also the procedures and conventions for administering and scoring tests must be standardized. Among the issues of relevance to this topic are the amount of pre-stress practice required to achieve stable performance on each test, and the development of performance metrics for each test which are representative of the categories in the performance degradation taxonomy. The research which has been conducted to date has not attempted to systematically examine these issues for all of the tests in the SPB. Generation of the information necessary for complete standardization of the test battery will require substantial effort in future phases of the research. As a result of the first year research activities, some preliminary data on these topics are available from the study of noise effects on tests from the prototype SPB.

The data from the research on the effects of noise on the performance categories of Reaction Time, Rate Control, and Attention constitute the second major product of the current research. The research examined the effects of patterned and randomly intermittent noise on Reaction Time, Rate Control, and Time-Sharing tasks. Although the data derived from this study are preliminary in nature, they provide considerable support for the approach which has been taken to the assessment of stress effects. These data also evidence the potential of the SPB for precisely detecting and diagnosing the effects of environmental stressors. The results which were obtained confirmed the need for broad based performance assessment by demonstrating that stressors can differentially affect various categories of performance. This was illustrated in the data by the presence of effects only for the randomly intermittent noise and by the selective appearance of these effects on Reaction Time

performance. The data further revealed that not all measures of performance are equally sensitive to the presence of stress effects. The research findings indicated that the effects of random noise were restricted to performance measures reflecting Category 4 (decreases in performance rate) from the performance degradation taxonomy. This finding emphasizes the fact that the effective diagnosis of stress effects does not depend exclusively on the development of a standardized battery of performance tests. Assessment of stress effects also requires the derivation of performance measures which are sufficiently sensitive and comprehensive to reveal the variety of ways in which stress affects the particular performances represented by tests in the SPB. The current research has been aimed primarily at the development of a standardized battery of tests. Therefore, research in subsequent phases of the project will be needed to derive the requisite performance metrics. Successful completion of both of these developmental efforts will result in a truly comprehensive instrument for evaluating stress effects.

The study of the effects of noise additionally disclosed the need for research on the time course of noise effects. The suggestion in the data, that certain noise stressors may have their maximum effects on some tests after a brief exposure and on others only after a longer period of time, has significant implications for the design of tests in the SPB. If it can be firmly established that noise effects become manifest in some cases only after extended performance under noise stress, versions of the performance tests compatible with repeated or extended performance measurement might have to be developed. However, before such a test development effort is undertaken, preliminary experimentation is required. It must be determined whether the delayed effects of noise can be observed only after prolonged performance under noise stress, or whether they can be detected following extended exposure to noise prior to the initiation of performance on a test. Only if the former is true, will the test development effort be required. The latter finding simply would entail an adjustment in the procedures for administration of performance tests and noise stressors.

An important test selection criterion employed in this research and recommended previously by Alluisi and Fulkerson (1964) is trainability. This criterion requires that asymptotic levels of performance be quickly attainable for each test during pre-stress administration. The data collected during the study of noise effects indicate that the three tests which were examined meet this criterion. In addition, since the Time Sharing task is probably the most difficult in the prototype SPB, it is possible to infer that stable performance on the remaining tests can be achieved through the use of relatively brief periods of practice.

The final conclusion derivable from the noise experiment is that physiological covariates may be necessary to assess the impact of stressors on experimental subjects. There is some indication in the analyses which have been conducted that inter-subject variability in reaction to stress may interfere with the detection of stress effects. If further analysis of these data or additional research can verify this observation, the measurement of physiological parameters may be necessary to insure that a given "stressor" is indeed stressful for each of the experimental subjects. Oken, et al. (1966) and Lacey (1967) have suggested several measures suitable for this purpose. Among these are heart rate, skin conductance, blood pressure, and EMG.

The third major product of this research effort was the formulation of a methodology for diagnostically categorizing the noise literature. The format for organizing this literature will eventually be based upon the integration of the categories from the performance degradation taxonomy with the categories of performance underlying the SPB. Since a final set of performance categories cannot be specified until a terminal form of the SPB is developed, the noise literature has been provisionally organized only with respect to the degradation categories. An examination of the classified literature highlighted the differential sensitivity of the degradation categories to noise stress, and underscored the need for developing performance metrics to represent each of the degradation categories. This classificational system will not only be of use

in organizing the noise literature but will be generally useful for categorizing and interpreting the effects of any stressor. It will also serve as a basis for comparing stress effects across different types of stressors and performance tasks. Application of this system to the stress literature will contribute to the development of a systematic body of knowledge concerning stress effects.

In general, it can be concluded that the research activities which have been completed have resulted in substantial progress toward the goal of developing a standardized test battery for the assessment of the effects of noise and other environmental stressors. These efforts also have provided concrete guidelines for further development of the test battery and its accompanying methodology.

REFERENCES CITED IN THE REPORT

Alluisi, E. A. Performance decrement in air mobility. In Report of the Tenth Annual Army Human Factors Research and Development Conference, U. S. Army Aviation Center, 1964. Pp. 213-216.

Alluisi, E. A. Methodology in the use of synthetic tasks to assess complex performance. Human Factors, 1967, 9(4), 375-384.

Alluisi, E., Chiles, W., & Hall, T. Combined effects of sleep loss and demanding work-rest schedules on crew performance. AMRL Technical Documentary Report No. 64-63, Aerospace Medical Research Laboratories, Wright-Patterson Air Force Base, Ohio, 1964.

Alluisi, E., Chiles, W., Hall, T., & Hawkes, G. Human group performance during confinement. AMRL Technical Documentary Report No. 63-87, Aerospace Medical Research Laboratories, Wright-Patterson Air Force Base, Ohio, 1963.

Alluisi, E. A. & Fulkerson, S. C. Behavioral effects of infectious disease: Semi-annual progress report, 1 March 1964-31 August 1964. Report No. PR-64-2, Performance Research Laboratory, University of Louisville, Louisville, Kentucky, 1964.

Alluisi, E. A., Thurmond, J. B., & Fulkerson, S. C. Behavioral effects of infectious diseases. Report No. PR-66-5, Performance Research Laboratory, University of Louisville, Louisville, Kentucky, 1966.

Atherley, G. R., Gibbons, S. L., & Powell, J. A. Moderate acoustic stimuli: The interrelation of subjective importance and certain physiological changes. Ergonomics, 1970, 13(5), 536-545.

Baker, W. J., Elkin, E. H., Van Cott, H. P., & Fleishman, E. A. Effects of drugs on human performance: The development of analytic methods and tests of basic human abilities. Report No. AIR-E-25-3/66-TR-2, American Institutes for Research, Washington, D. C., 1966.

Baker, W. J. & Theologus, G. C. The effects of caffeine on long-term visual monitoring. Report No. AIR-867-4/70-FR, American Institutes for Research, Washington, D. C., April 1970.

Bills, A. G. Blocking: A new principle of mental fatigue. American Journal of Psychology, 1931, 43, 230-245.

Bills, A. G. Some causal factors in mental blocking. Journal of Experimental Psychology, 1935, 18, 172-185.

REFERENCES (Cont.)

Bills, A. G. Blocking in mental fatigue and anoxemia compared. Journal of Experimental Psychology, 1937, 20, 437-452.

Biotechnology, Inc. Operating instructions: Perceptual-motor performance console Model 776. Arlington, Virginia, 1966.

Chambers, R. M. Operator performance in acceleration environments. In N. M. Burns, R. M. Chambers, & E. Hendler (Eds.), Unusual environments and human behavior. London: Glencoe, 1963. Pp. 193-319.

Duncanson, J. P. Learning and measured abilities. Journal of Educational Psychology, 1966, 57, 220-229.

Evans, W. O. & Consolazio, F. C. Effects of high altitude on performance of three different types of work. Perceptual and Motor Skills, 1967, 25, 41-50.

Finkelman, J. M. & Glass, D. C. A reappraisal of the relationship between noise and human performance by means of a subsidiary task measure. Experimental Publication System, American Psychological Association, 1969. Issue No. 3, Manuscript No. 090B.

Fleishman, E. A. The structure and measurement of physical fitness. New Jersey: Prentice-Hall, 1964.

Fleishman, E. A. & Hempel, W. E. A factor analysis of dexterity tests. Personnel Psychology, 1954, 7, 15-32.

French, J. W., Ekstrom, R. B., & Price, L. A. Manual for kit of reference tests for cognitive factors. Princeton, N. J.: Educational Testing Service, 1963.

Gorham, W. A. & Orr, D. B. Research on behavior impairment due to stress. Report No. AIR-B14-57-IR-64, American Institutes for Research, Washington, D. C., September 1957.

Gorham, W. A. & Orr, D. B. Research on behavior impairment due to stress: Experiments in impairment reduction. Report No. AIR-B14-58-FR-189, American Institutes for Research, Washington, D. C., November 1958.

Guilford, J. P. When not to factor analyze. Psychological Bulletin, 1952, 49, 26-37.

Guilford, J. P. The nature of human intelligence. New York: McGraw-Hill, 1967.

REFERENCES (Cont.)

Guilford, J. P., Berger, R. M., & Christensen, P. R. A factor-analytic study of planning. II. Administration of tests and analysis of results. Report No. 12, Psychology Laboratory, University of Southern California, Los Angeles, 1955.

Guilford, J. P. & Christensen, P. R. A factor-analytic study of verbal fluency. Report No. 17, Psychological Laboratory, University of Southern California, Los Angeles, 1956.

Guilford, J. P., Frick, J. W., Christensen, P. R., & Merrifield, P. R. A factor-analytic study of flexibility in thinking. Report No. 18, Psychological Laboratory, University of Southern California, Los Angeles, 1957.

Guilford, J. P. & Haepfner, R. Sixteen divergent-production abilities at the ninth-grade level. Multivariate Behavioral Research, 1966, 1(1), 43-66.

Guilford, J. P., Kettner, N. W., & Christensen, P. R. A factor-analytic investigation of the factor called general reasoning. Report No. 14, Psychological Laboratory, University of Southern California, Los Angeles, 1955.

Guilford, J. P., Kettner, N. W., & Christensen, P. R. A factor-analytic study across the domains of reasoning, creativity, and evaluation. II. Administration of tests and analysis of results. Report No. 16, Psychological Laboratory, University of Southern California, 1956.

Guilford, J. P., Wilson, R. C., & Christensen, P. R. A factor-analytic study of creative thinking. Report No. 8, Psychological Laboratory, University of Southern California, Los Angeles, 1952.

Kryter, K. D., Ward, W. D., Miller, J. D., & Eldredge, D. H. Hazardous exposure to intermittent and steady-state noise. Journal of the Acoustical Society of America, 1966, 39(3), 451-464.

Lacey, J. I. Somatic response patterning and stress: Some revisions of activation theory. In M. H. Appley & R. Trumbell (Eds.), Psychological stress. New York: Appleton-Century-Crofts, 1967.

Lemke, E. A., Klausmeier, H. J., & Harris, C. W. Relationship of selected cognitive abilities to concept attainment and information processing. Journal of Educational Psychology, 1967, 58(1), 27-35.

REFERENCES (Cont.)

McCann, P. H. The effects of ambient noise on vigilance performance. Human Factors, 1969, 11(3), 251-256.

McNemar, Q. The factors in factoring behavior. Psychometrika, 1961 16(4), 353-359.

Oken, D., Heath, H., Shipman, W., Goldstein, I., Grinker, R., & Fisch, J. The specificity of response to stressful stimuli. Archives of General Psychiatry, 1966, 15(6), 624-634.

Parker, J. F., Reilly, R. E., Dillon, R. F., Andrews, T. G., & Fleishman, E. A. Development of tests for measurement of primary perceptual-motor performance. Report No. NASA CR-335, National Aeronautics and Space Administration, Washington, D. C., December 1965.

Sanders, A. F. The influence of noise on two discrimination tasks. Ergonomics, 1961, 4(3), 253-258.

Sullivan, R., Warren, R., & Dabice, M. Minimal aversion thresholds for white noise: Adaptation. American Journal of Psychology, 1970, 83(4), 613-620.

Tune, G. S. Psychological effects of hypoxia: A review of certain literature from the period 1950 to 1963. Perceptual and Motor Skills, 1964, 19, 551-562.

Wyon, D. P. Studies of children under imposed noise and heat stress. Ergonomics, 1970, 13(5), 598-612.

APPENDIX A

References for the Categories of Performance in the Standardized Performance Battery

APPENDIX A

List of Factors used for the Standardized Performance Battery and their Related References

Factor	References
Psychomotor	
Control Precision	29, 32, 37, 38, 39, 87
Multilimb Coordination	32, 39, 87
Response Orientation	30, 31, 32, 39, 87
Reaction Time	29, 32, 38, 87
Speed of Arm Movement	29, 32, 34, 37, 38, 87
Rate Control	32, 38, 39
Manual Dexterity	27, 28, 34, 36, 39, 71, 72, 87
Finger Dexterity	27, 28, 34, 36, 71, 72, 87
Arm-Hand Steadiness	27, 28, 32, 33, 38, 71, 72, 87
Wrist-Finger Speed	27, 28, 34, 36, 52
Aiming	27, 28, 34, 36, 71, 72
Perceptual	
Flexibility of Closure	10, 11, 12, 40, 53, 62, 65, 86, 88, 93
Speed of Closure	10, 11, 35, 40, 44, 45, 46, 47, 51, 53, 58, 62, 65, 76, 86, 88, 93, 101
Length Estimation	2.6, 2.7, 2.10, 2.12, 44, 57, 92, 93, 105
Perceptual Speed	2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 2.10, 2.11, 2.12, 2.13, 2.14, 2.15, 6, 7, 10, 17, 19, 20, 23, 24, 25, 37, 39, 41, 42, 44, 45, 46, 47, 51, 52, 53, 57, 58, 60, 65, 66, 69, 80, 84, 85, 91, 92, 93, 95, 96, 97, 98, 99, 100, 102, 108, 110
Spatial Orientation	2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 2.10, 2.11, 2.12, 2.13, 2.14, 2.15, 3, 6, 15, 16, 19, 20, 41, 44, 45, 48, 49, 50, 51, 53, 67, 70, 84, 85, 88, 89, 97, 98, 99, 100, 103, 106
Spatial Scanning	1, 8, 39, 41, 53, 57, 90, 100

APPENDIX A (Cont.)

Factor	References
Perceptual (Cont.)	
Visualization	1, 2.1, 2.2, 2.5, 2.6, 2.7, 2.8, 2.9, 2.10, 2.11, 2.12, 2.13, 2.15, 5, 8, 24, 30, 37, 39, 41, 42, 44, 45, 46, 51, 53, 54, 57, 58, 62, 65, 66, 73, 78, 92, 93, 101, 102, 109, 110
Cognitive	
Associational Fluency	6, 13, 18, 41, 45, 53, 55, 58, 60, 64, 65, 78, 79, 82, 83, 100, 102
Expressional Fluency	13, 18, 41, 42, 53, 55, 60, 64, 95
Ideational Fluency	6, 8, 13, 41, 43, 54, 55, 58, 60, 61, 65, 74, 95
Word Fluency	1, 6, 7, 11, 12, 13, 18, 41, 45, 46, 53, 55, 63, 64, 65, 66, 67, 88, 95, 97, 99, 100, 102
Induction	1, 10, 11, 20, 41, 49, 58, 60, 68, 97, 98, 99, 100
Associative Memory	1, 2.13, 2.14, 7, 13, 17, 20, 21, 41, 48, 49, 53, 57, 66, 77, 92, 93, 95, 97, 98, 99, 100, 104, 107, 109
Mechanical Knowledge	1, 2.2, 2.3, 2.4, 2.5, 2.8, 2.9, 2.10, 2.11, 2.12, 7, 17, 19, 24, 30, 36, 38, 41, 42, 56, 70, 84, 89, 92, 108
Memory Span	3, 9, 41, 44, 53, 75, 76, 77, 92, 106
Number Facility	4, 8, 11, 12, 17, 37, 44, 46, 47, 51, 53, 54, 55, 58, 62, 65
Originality	8, 18, 26, 43, 53, 54, 55, 60, 61, 65, 78, 79, 83, 102
General Reasoning	2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 2.10, 2.11, 2.12, 2.13, 2.14, 2.15, 3, 4, 8, 10, 12, 13, 14, 16, 19, 20, 22, 39, 41, 43, 45, 49, 51, 53, 54, 55, 56, 58, 61, 62, 63, 64, 65, 68, 73, 79, 82, 83, 84, 85 88, 97, 98, 99, 101, 103, 104, 106, 109
Semantic Redefinition	53, 58, 65, 73, 102

APPENDIX A (Cont.)

Factor	References
Cognitive (Cont.)	
Syllogistic Reasoning	1, 8, 11, 51, 53, 54, 56, 58, 61, 62, 73, 83, 107
Sensitivity to Problems	8, 53, 54, 60, 65, 78, 79, 82, 102
Verbal Comprehension	2.2, 2.3, 2.4, 2.5, 2.8, 2.9, 2.10, 2.11, 2.12, 2.13, 2.14, 2.15, 6, 8, 11, 13, 14, 16, 17, 18, 19, 20, 22, 24, 41, 43, 44, 45, 47, 48, 49, 51, 53, 54, 56, 58, 60, 61, 65, 67, 68, 70, 74, 84, 89, 95, 96, 97, 98, 99, 100, 105, 106
Figural Adaptive Flexibility	8, 43, 53, 54, 61, 65, 78, 102
Semantic Spontaneous Flexibility	43, 53, 60, 61, 65, 102

APPENDIX A (Cont.)

References to Appendix A

1. Adkins, D. & Lyerly, S. Factor analysis of reasoning tests. Chapel Hill, N.C.: University of North Carolina Press, 1952.

2. Army Air Forces Aviation Psychology Program Research Reports, Printed Classification Tests. Report No. 5. J. P. Guilford & J. Lacey (Eds.). Washington, D.C.: U. S. Government Printing Office, 1947.

2.1 Army Air Forces Printed Classifications Tests, Carefulness Battery. Report No. 5. Pp. 686-695.

2.2 Army Air Forces Printed Classification Tests, December 1942 Classification Battery. Report No. 5. Pp. 798-821.

2.3 Army Air Forces Printed Classification Tests, July 1943 Classification Battery. Report No. 5. Pp. 798-821.

2.4 Army Air Forces Printed Classification Tests, November 1943 Classification Battery. Report No. 5. Pp. 798-821.

2.5 Army Air Forces Printed Classification Tests, September 1944 Classification Battery. Report No. 5. Pp. 798-821.

2.6 Army Air Forces Printed Classification Tests, Perceptual Battery I. Report No. 5. Pp. 408-418.

2.7 Army Air Forces Printed Classification Tests, Perceptual Battery II. Report No. 5. Pp. 408-418.

2.8 Army Air Forces Printed Classification Tests, Foresight and Planning Tests I. Report No. 5. Pp. 180-190.

2.9 Army Air Forces Printed Classification Tests, Foresight and Planning Tests II. Report No. 5. Pp. 180-190.

2.10 Army Air Forces Printed Classification Tests, Integration Tests. Report No. 5. Pp. 215-225.

2.11 Army Air Forces Printed Classification Tests, Judgment Tests. Report No. 5. Pp. 142-155.

2.12 Army Air Forces Printed Classification Tests, Mechanical Tests. Report No. 5. Pp. 333-339.

2.13 Army Air Forces Printed Classification Tests, Memory Tests I. Report No. 5. Pp. 261-268.

APPENDIX A (Cont.)

References

2.14 Army Air Forces Printed Classification Tests, Memory Tests II. Report No. 5. Pp. 261-268.

2.15 Army Air Forces Printed Classification Tests, Reasoning Tests. Report No. 5. Pp. 113-122.

3. Balinski, B. An analysis of the mental factors of various age groups from nine to sixty. Genetic Psychology Monographs, 1941, 23, 191-234.

4. Barakat, M. A factorial study of mathematical abilities. British Journal of Psychology, 1951, 9(3), 137-156.

5. Barratt, E. An analysis of verbal reports of solving spatial problems as an aid in defining spatial factors. Journal of Psychology, 1953, 36, 17-25.

6. Bechtoldt, H. Factorial investigation of the perceptual speed factor. American Psychologist, 1947, 2, 304-305.

7. Bechtoldt, H. Factor analysis of the Airman Classification Battery with civilian reference tests. USAF Human Resources Research Center Research Bulletin No. 53-59, 1953.

8. Berger, R., Guilford, J., & Christensen, P. A factor-analytic study of planning abilities. Psychological Monographs, 1957, 71, 1-31 (Whole No. 435).

9. Birren, J. A factor analysis of the Wechsler-Bellevue Scale given to an elderly population. Journal of Consulting Psychology, 1952, 16, 399-405.

10. Blakey, R. A factor analysis of a non-verbal reasoning test. Educational and Psychological Measurement, 1941, 189-198.

11. Botzum, W. A. A factorial study of the reasoning and closure factors. Psychometrika, 1951, 16, 361-386.

12. Canisia, M. Mathematical ability as related to reasoning and use of symbols. Educational and Psychological Measurement, 1962, 22(1), 105-127.

13. Carroll, J. B. A factor analysis of verbal abilities. Psychometrika, 1941, 6(5), 279-307.

14. Carroll, J. B. The factorial representation of mental ability and academic achievement. Educational and Psychological Measurement, 1943, 3, 307-332.

APPENDIX A (Cont.)

References

15. Chapman, R. L. The MacQuarrie test for mechanical ability. Psychometrika, 1948, 13, 175-179.

16. Chein, I. An experimental study of verbal, numerical, and spatial factors in mental organization. Psychological Record, 1939, 3, 71-94.

17. Christal, R. E. Factor analytic study of visual memory. Psychological Monographs, 1958, 72, 1-24 (Whole No. 466).

18. Christensen, P. R. & Guilford, J. P. An experimental study of verbal fluency factors. British Journal of Statistical Psychology, 1963, 16(1), 2-26.

19. Comrey, A. L. A factorial study of achievement in West Point courses. Educational and Psychological Measurement, 1949, 9, 193-209.

20. Coombs, C. H. A factorial study of number ability. Psychometrika, 1941, 6, 161-189.

21. Corter, H. M. Factor analysis of some reasoning tests. Psychological Monographs, 1952 (Whole No. 340).

22. Davidson, W. M. & Carroll, J. B. Speed and level components in time-limit scores: A factor analysis. Educational and Psychological Measurement, 1945, 5, 411-427.

23. Davis, P. A factor analysis of the Wechsler-Bellevue Scale. Educational and Psychological Measurement, 1956, 16, 127-146.

24. Dudek, F. J. The dependence of factorial composition of aptitude tests upon population differences among pilot trainees. I. The isolation of factors. Educational and Psychological Measurement, 1948, 8, 613-633.

25. Dysinger, D. W. Factorial study of speed of response in simple cancellation tasks. Iowa Academy of Science, 1959, 57, 373-378.

26. Fee, F. An alternative to Ward's factor analysis of Wallach and Kogan's "creativity" correlations. British Journal of Educational Psychology, 1968, 38(3), 319-321.

27. Fleishman, E. A. Testing for psychomotor abilities by means of apparatus tests. Psychological Bulletin, 1953, 50, 241-262.

APPENDIX A (Cont.)

References

28. Fleishman, E. A. A factorial study of psychomotor abilities. Research Bulletin AFPTRC-TR-54-15, May 1954. Air Force Personnel and Training Center, Lackland Air Force Base, Texas.

29. Fleishman, E. A. Dimensional analysis of psychomotor abilities. Journal of Experimental Psychology, 1954, 48, 437-454.

30. Fleishman, E. A. A comparative study of aptitude patterns in unskilled and skilled motor performances. Journal of Applied Psychology, 1957, 41, 263-272.

31. Fleishman, E. A. Apports de Binet aux tests psycho-moteurs et development interieur de ces techniques. Revue de Psychologie Applique, 1957, 7, 231-248.

32. Fleishman, E. A. Analysis of positioning movements and static reactions. Journal of Experimental Psychology, 1958, 55, 13-24.

33. Fleishman, E. A. A relationship between incentive motivation and ability level in psychomotor performance. Journal of Experimental Psychology, 1958, 56, 78-81.

34. Fleishman, E. A. & Ellison, G. D. A factor analysis of fine manipulation tests. Journal of Applied Psychology, 1962, 46(2), 96-105.

35. Fleishman, E. A. & Fruchter, B. Factor structure and predictability of successive stages in learning Morse code. Journal of Applied Psychology, 1960, 44, 97-101.

36. Fleishman, E. A. & Hempel, W. E. A factor analysis of dexterity tests. Personnel Psychology, 1954, 7, 15-32.

37. Fleishman, E. A. & Hempel, W. E. Changes in factor structure of a complex psychomotor test as a function of practice. Psychometrika, 1954, 19, 239-252.

38. Fleishman, E. A. & Hempel, W. E. The relation between abilities and improvement with practice in a visual discrimination reaction task. Journal of Experimental Psychology, 1955, 49, 301-310.

39. Fleishman, E. A. & Hempel, W. E. Factorial analysis of complex psychomotor performance and related skills. Journal of Applied Psychology, 1956, 40, 96-104.

APPENDIX A (Cont.)

References

40. Fleishman, E. A., Roberts, M. M., & Friedman, M. P. A factor analysis of aptitude and proficiency measures in radiotelegraphy. Journal of Applied Psychology, 1958, 42, 123-135.

41. French, J. W. The description of aptitude and achievement factors in terms of rotated factors. Psychometric Monographs, No. 5, 1951.

42. French, J. W. The factorial invariance of pure-factor tests. Journal of Educational Psychology, 1957, 48, 93-109.

43. Frick, J. W., Guilford, J. P., Christensen, P. R., & Merrifield, P. R. A factor-analytic study of flexibility in thinking. Educational and Psychological Measurement, 1959, 19(4), 469-496.

44. Fruchter, B. Factorial content of wrong-response scores of aptitude tests. Paper presented at the Western Psychological Association, San Francisco, June 1948.

45. Fruchter, B. The nature of verbal fluency. Educational and Psychological Measurement, 1948, 8,33-47.

46. Fruchter, B. A factor analysis of aptitude, achievement, and background variables. Paper presented at the meeting of the Western Psychological Association, Eugene, Oregon, June 1949.

47. Fruchter, B. Orthogonal and oblique solutions of a battery of aptitude, achievement, and background variables. Educational and Psychological Measurement, 1952, 12, 20-38.

48. Garrett, H. E. Differential mental traits. Psychological Record, 1938, 2, 259-298.

49. Goodman, C. H. A factorial analysis of Thurstone's sixteen primary mental abilities tests. Psychometrika, 1943, 8, 141-151.

50. Goodman, C. H. The MacQuarrie test for mechanical ability: II. Factor analysis. Journal of Applied Psychology, 1947, 31, 150-154.

51. Green, R. F., Guilford, J. P., Christensen, P. R., & Comrey, A. L. A factor-analytic study of reasoning abilities. Psychometrika, 1953, 18(2), 135-160.

52. Greene, E. B. An analysis of random and systematic changes with practice. Psychometrika, 1943, 8, 37-52.

APPENDIX A (Cont.)

References

53. Guilford, J. P. The nature of human intelligence. New York: McGraw-Hill, 1967.

54. Guilford, J. P., Berger, R. M., & Christensen, P. R. A factor-analytic study of planning. II. Administration of tests and analysis of results. Report No. 12, Psychological Laboratory, University of Southern California, Los Angeles, 1955.

55. Guilford, J. P. & Christensen, P. R. A factor-analytic study of verbal fluency. Report No. 17, Psychological Laboratory, University of Southern California, Los Angeles, 1956.

56. Guilford, J. P., Christensen, P. R., Kettner, N. W., Green, R. F., & Hertzka, A. F. A factor-analytic study of Navy Reasoning Tests with the Air Force Aircrew Classification Battery. Educational and Psychological Measurement, 1954, 14, 301-325.

57. Guilford, J. P., Frick, J. W., Christensen, P. R., & Merrifield, P. R. A factor-analytic study of flexibility in thinking, Report No. 18, Psychological Laboratory, University of Southern California, Los Angeles, 1957.

58. Guilford, J. P., Fruchter, B., & Zimmerman, W. S. Factor analysis of the Army Air Force Sheppard Field battery of experimental aptitude tests. Psychometrika, 1952, 17, 45-68.

59. Guilford, J. P., Hertzka, A. F., & Christensen, P. R. A factor-analytic study of evaluative abilities. II. Administration of tests and analysis of results. Report No. 9, Psychological Laboratory, University of Southern California, Los Angeles, 1953.

60. Guilford, J. P., Kettner, N. W., & Christensen, P. R. A factor-analytic investigation of the factor called general reasoning. Report No. 14, Psychological Laboratory, University of Southern California, Los Angeles, 1955.

61. Guilford, J. P., Kettner, N. W., & Christensen, P. R. A factor-analytic study across the domains of reasoning, creativity, and evaluation. II. Administration of tests and analysis of results. Report No. 16, Psychological Laboratory, University of Southern California, Los Angeles, 1956.

62. Guilford, J. P., Merrifield, P. R., Christensen, P. R., & Frick, J. W. An investigation of symbolic factors of cognition and convergent production. Report No. 23, Psychological Laboratory, University of Southern California, Los Angeles, 1960.

APPENDIX A (Cont.)

References

63. Guilford, J. P., Merrifield, P. R., Christensen, P. R., & Frick, J. W. Some new symbolic factors of cognition and convergent production. Educational and Psychological Measurement, 1961, 21, 515-542.

64. Guilford, J. P., Wilson, R. C., & Christensen, P. R. A factor-analytic study of creative thinking. Report No. 8, Psychological Laboratory, University of Southern California, Los Angeles, 1952.

65. Guthrie, G. M. Structure of abilities in a non-Western culture. Journal of Educational Psychology, 1963, 54, 94-103.

66. Hakkinen, S. & Ruoppila, I. The effect of performance time and retesting upon the factor structure of intelligence tests. Scandinavian Journal of Psychology, 1960, 1, 145-156.

67. Hall, W. E. & Robinson, F. P. An analytical approach to the study of reading skills. Journal of Educational Psychology, 1945, 36, 429-442.

68. Haney, R., Michael, W. B., & Jones, R. A. Identification of aptitude and achievement factors in the prediction of success in nursing training. Educational and Psychological Measurement, 1959, 19, 645-647.

69. Harrell, W. A factor analysis of mechanical ability tests. Psychometrika, 1940, 5, 17-33.

70. Hempel, W. E. & Fleishman, E. A. A factor analysis of physical proficiency and manipulative skill. Research Bulletin No. 54-34, Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas, 1954.

71. Hempel, W. E. & Fleishman, E. A. A factor analysis of physical proficiency and manipulative skill. Journal of Applied Psychology, 1955, 39, 12-16.

72. Hertzka, A. F., Guilford, J. P., Christensen, P. R., & Berger, R. M. A factor-analytic study of evaluative abilities. Educational and Psychological Measurement, 1954, 14, 581-597.

73. Johnson, D. M. & Reynolds, F. A factor analysis of verbal ability. Psychological Record, 1941, 4, 183-195.

74. Jones, L. V. A factor analysis of the Stanford-Binet at four age levels. Psychometrika, 1949, 14, 299-331.

APPENDIX A (Cont.)

References

75. Jones, L. V. Primary abilities in the Stanford-Binet, age 13. Journal of Genetic Psychology, 1954, 84, 125-147.

76. Kelley, H. P. A factor analysis of memory ability. (Ph.D. Thesis, Princeton University, 1954.) Research Bulletin 54-7, Educational Testing Service, Princeton, N. J., 1954.

77. Kettner, N. W., Guilford, J. P., & Christensen, P. R. The relation of certain thinking factors to training criteria in the U. S. Coast Guard Academy. Educational and Psychological Measurement, 1959, 19, 381-394.

78. Kettner, N. W., Guilford, J. P., & Christensen, P. R. A factor-analytic study across the domains of reasoning, creativity, and evaluation. Psychological Monographs, 1959 (Whole No. 479).

79. Kleemeier, R. W. & Dudek, F. J. A factorial investigation of flexibility. Educational and Psychological Measurement, 1950, 10, 107-118.

80. Langsam, R. S. A factorial analysis of reading ability. Journal of Experimental Education, 1941, 10, 57-63.

81. Merrifield, P. R., Guilford, J. P., Christensen, P. R., & Frick, J. W. A factor-analytic study of problem-solving abilities. Report No. 22, Psychological Laboratory, University of Southern California, Los Angeles, 1960.

82. Merrifield, P. R., Guilford, J. P., Christensen, P. R., & Frick, J. W. The role of intellectual factors in problem solving. Psychological Monographs, 1962, 76, No. 10 (Whole No. 529).

83. Michael, W. B. Factor analysis of tests and criteria: A comparative study of two AAF pilot populations. Psychological Monographs, 1951.

84. Michael, W. B., Zimmerman, W. S., & Guilford, J. P. An investigation of the nature of spatial relations and visualization factors in two high school samples. Educational and Psychological Measurement, 1951, 561-577.

85. Mooney, C. M. A factorial study of closure. Canadian Journal of Psychology, 1954, 8, 51-60.

86. Parker, J. F. & Fleishman, E. A. Ability factors and component performance measures as predictors of complex tracking behavior. Psychological Monographs, 1960, 70 (Whole No. 503).

References

87. Pemberton, C. The closure factors related to other cognitive processes. Psychometrika, 1952, 17, 267-288.

88. Peterson, D. A. A factor analysis of the new United States Basic Classification Test Battery. Washington, D. C.: U. S. Department of Commerce, 1946.

89. Remondion, C. A factorial analysis of the evaluation of scholastic compositions in the mother tongue. British Journal of Educational Psychology, 1959, 29, 242-251.

90. Rethlingshafer, D. The relation of tests of persistence to other measures of continuance of action. Journal of Abnormal and Social Psychology, 1942, 37, 71-82.

91. Roff, M. Personnel selection and classification procedures: Spatial tests; a factorial analysis. Project No. 21-29-002, U. S. Air Force School of Aviation Medicine, Randolph Field, Texas, 1951.

92. Roff, M. A factorial study of tests in the perceptual area. Psychometric Monographs, 1953, No. 8.

93. Sisk, H. L. A multiple factor analysis of mental abilities in the freshman engineering curriculum. Journal of Psychology, 1940, 9, 165-177.

94. Taylor, C. W. A factorial study of fluency in writing. Psychometrika, 1947, 12, 239-262.

95. Thornton, G. R. A factor analysis of tests designed to measure persistence. Psychological Monographs, 1939, 51, No. 229.

96. Thurstone, L. L. The perceptual factor. Psychometrika, 1938, 3, 1-17.

97. Thurstone, L. L. Experimental study of simple structure. Psychometrika, 1940, 5, 153-168.

98. Thurstone, L. L. Primary mental abilities. Psychometric Monographs, 1942, No. 1.

99. Thurstone, L. L. & Thurstone, T. G. Factorial studies of intelligence, Psychometric Monographs, 1942, No. 2.

100. Thurstone, L. L. A factorial study of perception. Psychological Monographs, 1944, No. 4.

APPENDIX A (Cont.)

References

101. Wilson, R. C., Guilford, J. P., Christensen, P. R., & Lewis, D. J. A factor-analytic study of creative-thinking abilities. Psychometrika, 1954, 19, 297-311.

102. Wittenborn, J. R. Factorial equations for tests of attention. Psychometrika, 1943, 8, 19-35.

103. Wittenborn, J. R., & Larsen, R. P. A factorial study of achievement in college German. Journal of Educational Psychology, 1944, 35, 39-48.

104. Woodrow, H. The relation between abilities and improvement with practice. Journal of Educational Psychology, 1938, 29, 215-230.

105. Woodrow, H. The common factors in fifty-two mental tests. Psychometrika, 1939, 4, 99-108.

106. Wrigley, C., Saunders, D. R., & Neuhaus, J. Application of the quartimax method of rotation to Thurstone's primary mental abilities study. Psychometrika, 1958, 23, 151-170.

107. Zachert, V., & Friedman, G. The stability of the factorial pattern of aircrew classification tests in four analyses. Psychometrika, 1953, 18, 219-224.

108. Zimmerman, W. S. A revised orthogonal rotational solution for Thurstone's original Primary Mental Abilities Test Battery, Psychometrika, 1953, 18, 77-93.

109. Zimmerman, W. S. The influence of item complexity upon the factor composition of a spatial visualization test. Educational and Psychological Measurement, 1954, 14, 106-119.

APPENDIX B

Diagnostic Classification of the Noise Literature

APPENDIX B

Provisional Definitions for the Performance Degradation Categories*

Category 1. Changes in Performance Distribution over Time

1.1 Lapses refer to irregularities in performance level over time caused by momentary lapses in attention. The concept is equivalent to that of response blocking and reflects distributions of performance characterized by periods of good performance cyclically interspersed with periods of poor performance.

1.2 Queuing represents a delay of response during peak loads followed by a compensatory acceleration during lulls. While similar to the concept of Lapses, Queuing implies a more deliberate delay of response than Lapses.

1.3 Error and/or escape from the performance task refer to rapid and prolonged increases in error amplitude and error frequency distribution. In some situations, performance may cease completely as though through drowsiness or refusal to perform. This subcategory is characterized by prolonged rather than cyclical interruptions of performance.

Category 2. Changes in Performance Distribution among Tasks, Among Task Components, or Within Task Components

2.1 Falling off in proficiency on parts of a task represents a selective reduction in performance on some task components or elements. Performance is reduced on some portions of the task while it is maintained on others.

2.2 Omissions of portions of tasks refer to the cessation rather than the reduction of performance on a task component.

2.3 Filtering represents a systematic filtering or omission of certain categories of stimulation or response according to some subjective priority scheme. The operator appears to be able to continue performance on all of the parts of the task but to ignore certain aspects of task performance.

*Adapted from Chambers (1963)

2.4 Fixation-block confusion is a behavioral syndrome characteristic of a narrowed attention associated with excessive concentration on one portion of the task. Concentration on a particular task element can reach the point where already defined stimuli regarding the situation as a whole are ignored. Typically, the subject fixates on one portion of the task, realizes that he has fixated and momentarily freezes or blocks. The block in performance is followed by a brief state of confusion. This may occur, for instance, in driving.

Category 3. Increases in Performance Rate

3.1 Performance oscillation refers to a high amplitude oscillation in operator response. It is found in response rather than error records. For example, on a tracking task it may not appear in integrated error records since the operator can move his control stick with an excessive rhythm and still maintain effective error compensation.

3.2 Inadvertent control inputs refer to false alarm responses in which an operator inserts a response when none is required.

3.3 Sudden changes in rate or frequency of performance refer to the sudden initiation or increase in rate of performance. The performance may be nonessential to the task.

Category 4. Decreases in Performance Rate

4.1 Changes in phasing and/or timing of task components represents a rhythmic slowing of reaction to serial events. For example, it can appear as a consistent lag in response to error deviations in a tracking task.

4.2 Response lags are represented by increased response latencies to discrimination stimuli.

Category 5. Changes in Performance Accuracy

5.1 Increases in error amplitude primarily refer to increases in absolute integrated error in tracking tasks.

5.2 Failure to detect refers to performance degradation in "go - no go" monitoring situations resulting from a failure to detect and respond to changes in the stimuli field.

5.3 <u>Error in returning, integrating, steering, and processing information</u>. No adequate definition for this subcategory currently exists.

5.4 <u>Perceived disintegration</u> of the perceptual field represents a narrowing of the perceptual field so that reduced attention is paid to the more peripheral indicators.

5.5 <u>Dissociation</u> of corrective response from the appropriate control unit reflects the situation in which an operator, in spite of having accurately detected the need for a particular corrective reaction, produces an inappropriate response.

5.6 <u>Stereotyping</u> of responses and movements refers to a loss of flexibility of set, and disjunction of discrimination and response. The subject performs as if all stimuli were equivalent, producing similar responses regardless of the stimulus situation.

5.7 <u>Approximations</u> refers to a degradation in performance in which a subject's performance may become less precise yet still within some specified tolerance limits. Responses become minimally adequate to meet the required criterion of proficiency.

APPENDIX B

Performance Taxonomy and Experimental Requirements for Redistribution of Performance Across Time

Item	NASA Reference	Noise Literature Reference	Dependent Measure	Task, Instrumentation and Problems
1.1 "Lapses"	Chambers, 1963, p. 274. Top pitch error reveals alternating periods of error and no error.	Fraser (1953) using var. of RT and Sanders (1961) using variance of difference scores for successive 1-min. periods showed stressor effects. Significant variables in Sanders: Task complexity (2.0 bit task more subject to stressor than a 1.5 bit task);Duration of exposure (stressor acted during second half of 30 min. sessions; Noise quality (randomly varying frequency vs. steady white noise). Plutchik (1961) showed increased variance of error score on a mirror tracing task under pulsed noise. Hack <u>et al</u>. (1965) using average derivatives of tracking error as measures of variability showed increased instability of performance under randomly pulsed noise of under 90 dB.	Variance, frequency of lapses, or derivatives of tracking error.	Any of the PEMCON tasks can provide needed data. PEMCON currently limited to recording 1-min. blocks of data. Some re-instrumentation needed to get finer resolution, e.g., 6 sec. blocks, for tracking. Problem is to keep the 1-min. trials but improve recording resolution. Need hard copy recording, i.e., event timer plus graphic level capability for tracking data. Alternative: employ operational amplifiers to perform continuous differentiation of tracking error curves.

APPENDIX B (Cont.)

Item	NASA Reference	Noise Literature Reference	Dependent Measure	Task, Instrumentation and Problems
1.2 "Queuing"			Mean: The experimental question is, "What happens on task A when the information load on task B is periodically increased, and then returned to "normal". Same question can be posed for sets of components and sets of elements.	Either tracking in combination with meter reading or RT, or tracking alone with load on one dimension varying. Could use stepping switch with set of fixed pots to change speed of motor controlling mechanical function generator. Could also use combinations of non-tracking tasks, i.e., finger tapping vs. meter reading where discrimininability of meter pairs is the dependent measure. Voltage to the integrators supplying the subject meters would have to be programmed, to vary systematically.
1.3 "Escape"				This item requires evidence that operator has been indifferent, i.e., that his responding is approaching or has reached chance levels. An appropriate dependent measure would be frequency of correct responses analyzed for random level. The experimental question might be,

APPENDIX B (Cont.)

Item	NASA Reference	Noise Literature Reference	Dependent Measure	Task, Instrumentation and Problems
1.3 (Cont.) "Escape"				"How long an exposure duration is required before random responding is initiated?" Possible task is the dual meter discrimination task, i.e. where subject determines whether meters do or do not have identical readings.
2.1 "Falling Off"	Chambers, 1963, p. 277. Control of yaw broke down under acceleration stress, but control of 3 concomittant dimensions suffered no loss.	Finkleman & Glass (1969). Under random noise (80 dB) a central tracking task was unimpaired. A subsidiary memory task, however, was impaired. Boggs & Simon (1968) Under randomly presented 5-sec. bursts of 92 dB noise showed greater impairment of a secondary auditory monitoring task when paired with a more complex CRT task than when paired with a less complex CRT task. CRT task unimpaired.	Either variance or mean. Analysis by frequency for tracking performance, or analysis of tracking error via differentiation.	Combination of tracking and monitoring. Might even push to 3 tasks, i.e. meter discrimination and RT, plus tracking with comparative analysis of x and y diemnsions. Linear frequency analysis to determine possible redistribution from fine to gross frequencies. Emphasize the variance measure. Special problem: to do frequency analysis we would need unintegrated data wave forms for signal input. This adds requirement for additional 4 graphic channels.
2.2 "Ommission"				
2.3 "Filtering"				
2.4 "Fixation block-confusion"				

APPENDIX B (Cont.)

Item	NASA Reference	Noise Literature Reference	Dependent Measure	Task, Instrumentation and Problems
3.1 "Performance oscillations"	Chambers, 1963, p. 275-276. Data shows marked build-up in amplitude of oscillatory responses to pitch, yaw, & roll deviations, during period of increasing G.		Amplitude of oscillatory responding. More specific criteria will have to be worked out. Chambers examples on visual inspection of raw data and to some extent, subjective decisions about the presence of "performance oscillations."	Compensatory tracking task is the most obvious way of getting at this item. We need, however, raw performance data, since the error data may mask excessive reactions, i.e., if compensatory oscillations are in effect, then ipso facto, we will not see evidence of performance decrement in the error data.
3.2 "Inadvertent control inputs" (false alarms)	Chambers 1963, p. 278, fig. 23 shows high amplitude		Ratio of output to input amplitude, i.e., response amplitude	This item would seem to require that the operator be fed small error inputs, requiring little or no effort by operator. Evidence for inadvertency would be seen in increasing response to error ratio as exposure to the stressor continued.

APPENDIX B (Cont.)

Item	NASA Reference	Noise Literature Reference	Dependent Measure	Task, Instrumentation and Problems
3.2 (Cont.)	control responses in reaction to zero deviations in roll, pitch, & yaw.		divided by error amplitude. For surveillance tasks the variable would be number of false alarms.	
3.3 "Sudden changes in rate or frequency"	Chambers 1963, p. 279. Cites data of subject required to press button at fixed rate of 60 per min. Rate increased to 250 per min. under 10 G.	McGuigan & Rodier (1968); auditory presentation of prose during reading led to a greater amplitude of covert oral behavior than occurred during silence, but white noise did not have this effect. Weinstein & Makenzie (1966). Under a 100 dB white noise manual performance on a block manipulation test increased.	Response frequency	Possible tasks: Compensatory tracking analyzed for operator frequency response via linear frequency theory. Other PEMCON tasks might be appropriate, e.g., hand steadiness test would provide response frequency data; Tapping test is another possibility. Variation on this would be paced tapping (See NASA reference). The tracking task would require raw amplitude vs. time curves for both the error input and operator responses.

APPENDIX B (Cont.)

Item	NASA Reference	Noise Literature Reference	Dependent Measure	Task, Instrumentation and Problems
4.1 "Changes in phasing..."	Chambers (1963), p. 274, lower part of Fig. 19 shows a systemmatic, cyclical error which essentially amounts to reduced reaction time for a continuous task.		Max. correlation on cross correlation function.	Requires that we correlate the raw curves for "vehicle" error and operator response as a function of phase shift between the two curves.
4.2 "Response lags"		A number of studies provide evidence that noise stressors can reduce response time, i.e. increase response lags. These include studies of monitoring, tracking, & arithmetic computation: Broadbent (1958); Grimaldi (1958); Loeb, *et al*. (1958); Teichner (1963); Reitter (1963); Wilkinson (1963). Most recent study by Wyon (1970) demonstrated reduced rate of performance on a numerical inspection task under 55-78 dB pulsed noise.	Reaction time or time to completion.	Possible PEMCON tasks: Meter discrimination and Audio-Visual RT task.

APPENDIX B (Cont.)

Item	NASA Reference	Noise Literature Reference	Dependent Measure	Task, Instrumentation and Problems
5.1 "Increases in error amplitude"	Chambers, (1963), p. 287 shows increase in integrated error as G level goes from +4 to +8.	A number of studies with tracking, arithmetic computation, and recall have shown that noise stressors can act to increase error amplitude in performance. These include: Burrows (1960); Dornic (1967); Grimaldi (1958); Hsia (1968); Mech (1953); Murray (1965); Wilkinson (1963); McCann (1969); and most recently, Eschenbrenner (1970).	Integrated absolute error. For the nontracking tasks, number of incorrect responses.	We already have capability for integrating tracking error on PEMCON. We need hard copy records, especially for smaller time periods than the 1-min. periods now programmed.
5.2 "Failures to detect.."	Chambers, (1963), p. 278. Doesn't give dat, but mentions failure to detect information on peripheral displays under high G	Mirabella & Goldstein's review (1967) concluded that noise stressors reduce detection probability under time sharing conditions, i.e., where multiple displays must be monitored for possible signals.	Number of critical events undetected.	At the present time there is no conventional detection task on PEMCON. We need to devise one, particularly since Chambers has specifically pointed to detection as source of performance decrement. One approach is to construct side panels with meter-type or warning light displays. Interesting aspect to this problem is need for "projective" detection, i.e., operator is told to respond with particular set of procedures when meter reading reaches a specified level. Much of space mission requirements depends on this kind of "projective" detection, i.e. "if abort light comes on and you are above Y thousand ft., do this; if below do that. Could get at this possibly via response orientation task, i.e. have S control vehicle and respons to intermittent RO signals. Measure accuracy and reaction time.

APPENDIX B (Cont.)

Item	NASA Reference	Noise Literature Reference	Dependent Measure	Task, Instrumentation and Problems
5.3 "Errors in returning, integrating.."	Chambers discusses this on p. 279. He may be referring to the ability to put together several sources of information which are either contiguous or sequential to arrive at a decision.		Number of correct decisions based on multiple data sources.	Again PEMCON is deficient. There is no task that requires coordination of multiple data sources to arrive at action decision, except two-hand coordinator. But motor coordination is not what Chambers' had in mind here. Possible task to get at this would be meter reading followed by RO where the coding scheme changes depending on meter position. This would roughly reflect situation in which abort actions are taken on the basis of altitude information.
5.4 "Perceived disintegration"				

APPENDIX B (Cont.)

Item	NASA Reference	Noise Literature Reference	Dependent Measure	Task, Instrumentation and Problems
5.5 "Dissociation"			Either per cent correct or bits, derived from a stimulus-response contingency matrix.	Response orientation task is one possible means for tapping this element. Could also add a panel to PEMCON with some configuration of lights and corresponding response devices.
5.6 "Sterotyping"				
5.7 "Approximations"				

APPENDIX B (Cont.)

References to Appendix B

Boggs, D., & Simon, J. Differential effect of noise on tasks of varying complexity. Journal of Applied Psychology, 1968, 52(2), 148-153.

Broadbent, D. Effect of noise on an "intellectual" task. Journal of the Acoustical Society of America, 1958, 30(9), 824-827.

Burrows, A. Acoustic noise, an informational definition. Human Factors, 1960, 2(3), 163-168.

Chambers, R. M. Operator performance in acceleration environments. In N. M. Burns, R. M. Chambers, & E. Hendler (Eds.), Unusual environments and human behavior. London: Glencoe, 1963, pp. 193-319.

Dornic, S. Effect of a specific noise on visual and auditory memory span. Scandinavian Journal of Psychology, 1967, 8, 155-160.

Eschenbrenner, A. Effects of intermittent noise on the performance of a complex psychomotor task. Human Factors, 1971, 13(1), 59-63.

Finkelman, J. M., & Glass, D. C. A reappraisal of the relationship between noise and human performance by means of a subsidiary task measure. Experimental publication system, American Psychological Association, 1969, Issue No. 3, Manuscript No. 090B.

Fraser, D. C. Recent experimental work in the study of fatigue. Occupational Psychology, 1958, 32, 258-263.

Grimaldi, J. Sensori-motor performance under varying noise conditions. Ergonomics, 1958, 2(1), 34-43.

Hack, J., Robinson, H., & Lathrop, R. Auditory distraction and compensatory tracking. Perceptual and Motor Skills, 1965, 20, 228-230.

Hsia, H. Effects of noise and difficulty level of input information in auditory, visual, and audiovisual information processing. Perceptual and Motor Skills, 1968, 26, 99-105.

Loeb, M., & Jeantheau, G. The influence of noxious environmental stimuli on vigilance. Journal of Applied Psychology, 1958, 42, 1, 47-49.

APPENDIX B (Cont.)

References

McGuigan, F. & Rodier, W. Effects of auditory stimulation on covert oral behavior during silent reading. Journal of Experimental Psychology, 1968, 76 (4), 649-655.

Mech, E. Factors influencing routine performance under noise. I. The influence of "set". Journal of Psychology, 1953, 35, 283-298.

Mirabella, A. & Goldstein, D. The effects of ambient noise upon signal detection. Journal of the Human Factors Society, 1967, 9(3), 277-284.

Murray, D. The effect of white noise upon the recall of vocalized lists. Canadian Journal of Psychology, 1965, 19(4), 333-345.

Plutchik, R. Effect of high intensity intermittent sound on compensatory tracking and mirror tracing. Perceptual and Motor Skills, 1961, 12, 187-194.

Reiter, H. Effects of noise on discrimination reaction time. Perceptual and Motor Skills, 1963, 17, 418.

Sanders, A. F. The influence of noise on two discrimination tasks. Ergonomics, 1961, 4(3), 253-258.

Teichner, W., Arees, E. & Reilly, R. Noise and human performance: A psychophysiological approach. Ergonomics, 1963, 6(1), 83-97.

Weinstein, A. & Mackenzie, R. Manual performance and arousal. Perceptual and Motor Skills, 1966, 22, 498.

Wilkinson, R. Interaction of noise with knowledge of results and sleep deprivation. Journal of Experimental Psychology, 1963, 66(4), 332-337.

Wyon, D. P. Studies of children under imposed noise and heat stress. Ergonomics, 1970, 13(5), 598-612.

APPENDIX C

Tables and Figures for the Study of the Effects of Noise

TABLE 5

Covariates and Unadjusted Mean Reciprocal Response Latencies in the Time Sharing and Reaction Time Tasks

Task	Noise Level	Covariate	Session 2				Session 3			
			Trial 1	Trial 2	Trial 3	Trial 4	Trial 1	Trial 2	Trial 3	Trial 4
TSRT*	0	2.174	2.237	2.177	2.159	2.162	2.315	2.261	2.242	2.287
	85	2.149	2.209	2.066	2.029	2.020	2.237	2.255	2.238	2.197
RT*	0	3.229	3.417	3.422	3.401	3.284	3.473	3.407	3.347	3.285
	85	3.243	3.138	3.147	3.215	3.093	3.308	3.319	3.346	3.319

*Time Sharing (reaction time)
**Reaction Time

TABLE 6

Covariates and Unadjusted Mean Frequencies of Response Blocks in the Time Sharing and Reaction Time Tasks

Task	Noise Level	Covariate	Session 2				Session 3			
			Trial 1	Trial 2	Trial 3	Trial 4	Trial 1	Trial 2	Trial 3	Trial 4
TSRT*	0 dB	7.300	3.100	3.500	3.800	3.300	1.700	2.800	2.900	2.800
	85 dB	6.900	3.900	4.700	5.300	5.100	2.700	2.800	2.500	3.300
RT**	0 dB	4.200	1.000	0.900	1.400	2.000	1.500	0.900	1.200	1.900
	85 dB	2.700	0.900	1.100	1.000	1.700	1.400	0.600	1.400	1.000

*Time Sharing (reaction time)
**Reaction Time

TABLE 7

Covariates and Unadjusted Mean Duration of Response Blocks in the Time Sharing and Reaction Time Tasks

Task	Noise Level	Covariate	Session 2				Session 3			
			Trial 1	Trial 2	Trial 3	Trial 4	Trial 1	Trial 2	Trial 3	Trial 4
TSRT*	0 dB	1.396	1.167	1.582	1.971	1.611	1.045	1.766	0.916	1.803
	85 dB	1.808	1.232	1.251	1.497	2.593	1.026	2.254	1.142	1.616
RT**	0 dB	0.796	0.456	0.395	0.316	0.688	0.528	0.303	0.311	0.574
	85 dB	0.576	0.416	0.339	0.444	0.471	0.378	0.239	0.481	0.340

*Time Sharing (reaction time)
**Reaction Time

TABLE 8

Covariates and Unadjusted Mean Reaction Times in the Time Sharing and Reaction Time Tasks

Task	Noise Level	Covariate	Session 2				Session 3			
			Trial 1	Trial 2	Trial 3	Trial 4	Trial 1	Trial 2	Trial 3	Trial 4
TSRT*	0 dB	0.454	0.443	0.454	0.458	0.457	0.438	0.444	0.448	0.439
	85 dB	0.459	0.445	0.473	0.472	0.476	0.446	0.445	0.448	0.456
RT**	0 dB	0.309	0.298	0.297	0.297	0.306	0.296	0.300	0.303	0.309
	85 dB	0.310	0.323	0.319	0.314	0.322	0.303	0.306	0.300	0.304

*Time Sharing (reaction time)
**Reaction Time

TABLE 9

Covariates and Unadjusted Mean Integrated Error on Both the X and Y Axes in the Time Sharing and Rate Control Tasks

Task	Noise Level	Covariate	Session 2				Session 3			
			Trial 1	Trial 2	Trial 3	Trial 4	Trial 1	Trial 2	Trial 3	Trial 4
TSRC*	0 dB	12.575	9.910	9.840	10.040	9.420	7.550	8.500	9.650	9.160
	85 dB	12.760	11.040	11.090	11.410	11.470	8.300	8.780	10.120	9.790
RC**	0 dB	10.635	8.190	8.230	8.930	7.510	6.290	6.400	8.310	6.870
	85 dB	10.440	7.710	8.150	8.910	8.190	6.680	7.490	7.380	7.950

*Time Sharing (rate control)
**Rate Control

TABLE 10

Covariates and Unadjusted Mean Integrated Error on the X - Axis in the Time Sharing and Rate Control Tasks

Task	Noise Level	Covariate	Session 2				Session 3			
			Trial 1	Trial 2	Trial 3	Trial 4	Trial 1	Trial 2	Trial 3	Trial 4
TSRC*	0 dB	11.390	8.900	8.540	9.000	8.200	6.520	7.400	8.240	8.120
	85 dB	11.860	10.500	10.340	10.700	10.620	7.520	7.900	9.360	9.440
RC**	0 dB	10.070	8.080	7.520	8.580	6.940	5.980	6.080	7.880	6.620
	85 dB	10.190	7.140	7.980	8.820	7.720	6.200	6.880	6.740	7.360

*Time Sharing (rate control)
**Rate Control

TABLE 11

Covariates and Unadjusted Mean Integrated Error on the Y - Axis in the Time Sharing and Rate Control Tasks

Task	Noise Level	Covariate	Session 2				Session 3			
			Trial 1	Trial 2	Trial 3	Trial 4	Trial 1	Trial 2	Trial 3	Trial 4
TSRC*	0 dB	13.760	10.920	11.140	11.080	10.640	8.580	9.600	11.060	10.200
	85 dB	13.660	11.580	11.840	12.120	12.320	9.080	9.660	10.880	10.140
RC**	0 dB	11.200	8.300	8.940	9.280	8.080	6.600	6.720	8.740	7.120
	85 dB	10.690	8.280	8.320	9.000	8.660	7.160	8.100	8.020	8.540

*Time Sharing (rate control)
**Rate Control

TABLE 12

Analyses of Variance and Covariance on Mean Transformed Response Latencies in Time Sharing Task

Source of Variation	Sum of Squares	*df*	Mean Square	*F* Ratio	*p* Value
Analysis of Variance					
Between Subjects					
A (Noise)	0.222	1	0.222	0.344	0.571
Error	11.637	18	0.646		
Within Subjects					
B (Session)	0.596	1	0.596	6.638	0.018
AB	0.031	1	0.031	0.344	0.571
Error	1.616	18	0.090		
C (Trial)	0.192	3	0.064	4.069	0.011
AC	0.020	3	0.007	0.430	0.736
Error	0.847	54	0.016		
BC	0.061	3	0.020	1.209	0.315
ABC	0.049	3	0.016	0.963	0.581
Error	0.910	54	0.017		
Analysis of Covariance					
Between Subjects					
A	0.125	1	0.125	0.507	0.508
Error	4.178	17	0.246		

TABLE 13

Analyses of Variance and Covariance on Mean Transformed Response Latencies in the Reaction Time Task

Source of Variation	Sum of Squares	df	Mean Square	F Ratio	p Value
			Analysis of Variance		
Between Subjects					
A (Noise)	0.841	1	0.841	1.097	0.310
Error	13.793	18	0.766		
Within Subjects					
B (Session)	0.304	1	0.304	8.851	0.008
AB	0.311	1	0.311	9.057	0.007
Error	0.618	18	0.034		
C (Trial)	0.226	3	0.075	3.449	0.022
AC	0.136	3	0.045	2.069	0.114
Error	1.181	54	0.022		
BC	0.031	3	0.010	0.554	0.652
ABC	0.019	3	0.006	0.331	0.805
Error	1.008	54	0.019		
			Analysis of Covariance		
Between Subjects					
A	0.953	1	0.953	2.905	0.103
Error	5.557	17	0.328		

TABLE 14

Analyses of Variance and Covariance on Mean Reaction Time in the Time Sharing Task

	Analysis of Variance				
Source of Variation	Sum of Squares	df	Mean Square	F Ratio	p Value
Between Subjects					
A (Noise)	0.004	1	0.004	0.218	0.650
Error	0.361	18	0.020		
Within Subjects					
B (Session)	0.009	1	0.009	2.040	0.168
AB	0.0003	1	0.0003	0.082	0.769
Error	0.075	18	0.004		
C (Trial)	0.006	3	0.002	5.660	0.002
AC	0.0008	3	0.0003	0.855	0.528
Error	0.018	54	0.0003		
BC	0.002	3	0.0006	1.313	0.279
ABC	0.001	3	0.0004	0.833	0.516
Error	0.024	54	0.0004		
	Analysis of Covariance				
Between Subjects					
A	0.002	1	0.002	0.306	0.593
Error	0.107	17	0.006		

TABLE 15

Analyses of Variance and Covariance on Mean Reaction Time in the Reaction Time Task

Source of Variation	Sum of Squares	df	Mean Square	F Ratio	p Value
		Analysis of Variance			
Between Subjects					
A (Noise)	0.005	1	0.005	0.734	0.593
Error	0.117	18	0.006		
Within Subjects					
B (Session)	0.002	1	0.002	5.347	0.031
AB	0.004	1	0.004	9.887	0.006
Error	0.006	18	0.0004		
C (Trial)	0.001	3	0.0004	2.422	0.075
AC	0.001	3	0.0002	1.575	0.205
Error	0.008	54	0.0002		
BC	0.0002	3	0.00007	0.566	0.644
ABC	0.00006	3	0.00002	0.151	0.925
Error	0.007	54	0.0001		
		Analysis of Covariance			
Between Subjects					
A	0.005	1	0.005	1.838	0.190
Error	0.042	17	0.002		

TABLE 16

Analyses of Variance and Covariance on Frequency of Response Blocks in the Time Sharing Task

Source of Variation	Sum of Squares	*df*	Mean Square	*F* Ratio	*p* Value
			Analysis of Variance		
Between Subjects					
A (Noise)	25.600	1	25.600	0.694	0.579
Error	663.875	18	36.882		
Within Subjects					
B (Session)	78.400	1	78.400	10.295	0.005
AB	11.025	1	11.025	1.448	0.243
Error	137.075	18	7.615		
C (Trial)	16.225	3	5.408	1.047	0.380
AC	2.350	3	0.783	0.152	0.925
Error	278.926	54	5.165		
BC	2.849	3	0.950	0.329	0.806
ABC	5.925	3	1.975	0.685	0.568
Error	155.725	54	2.884		
			Analysis of Covariance		
Between Subjects					
A	34.190	1	34.190	1.350	0.261
Error	430.537	17	25.326		

TABLE 17

Analyses of Variance and Covariance on Frequency of Response Blocks in the Reaction Time Task

Source of Variation	Sum of Squares	*df*	Mean Square	*F* Ratio	*p* Value
Analysis of Variance					
Between Subjects					
A (Noise)	1.806	1	1.806	0.217	0.651
Error	149.563	18	8.309		
Within Subjects					
B (Session)	0.006	1	0.006	0.008	0.892
AB	0.156	1	0.156	0.191	0.670
Error	14.712	18	0.817		
C (Trial)	12.119	3	4.040	5.524	0.003
AC	2.019	3	0.673	0.920	0.561
Error	39.488	54	0.731		
BC	4.819	3	1.606	1.787	0.159
ABC	2.269	3	0.756	0.841	0.520
Error	48.538	54	0.899		
Analysis of Covariance					
Between Subjects					
A	0.033	1	0.033	0.010	0.884
Error	57.729	17			

TABLE 18

Analyses of Variance and Covariance on Mean Duration of Response Blocks in the Time Sharing Task

Source of Variation	Sum of Squares	df	Mean Square	F Ratio	p Value
		Analysis of Variance			
Between Subjects					
A (Noise)	0.354	1	0.354	0.174	0.683
Error	36.624	18	2.035		
Within Subjects					
B (Session)	1.116	1	1.116	0.552	0.527
AB	0.043	1	0.043	0.021	0.857
Error	36.382	18	2.021		
C (Trial)	14.688	3	4.896	1.801	0.157
AC	1.446	3	0.482	0.177	0.909
Error	146.789	54	2.718		
BC	9.174	3	3.058	1.147	0.338
ABC	6.290	3	2.097	0.787	0.509
Error	143.9327	54	2.665		
		Analysis of Covariance			
Between Subjects					
A	0.096	1	0.096	0.458	0.514
Error	35.538	17	2.090		

TABLE 19

Analyses of Variance and Covariance on Mean Duration of Response Blocks in the Reaction Time Task

Source of Variation	Sum of Squares	df	Mean Square	F Ratio	p Value
Analysis of Variance					
Between Subjects					
A (Noise)	0.133	1	0.133	0.343	0.572
Error	6.992	18	0.388		
Within Subjects					
B (Session)	0.086	1	0.086	0.959	0.658
AB	0.006	1	0.006	0.062	0.793
Error	1.606	18	0.089		
C (Trial)	0.860	3	0.287	3.174	0.031
AC	0.722	3	0.241	2.664	0.056
Error	4.880	54	0.090		
BC	0.162	3	0.054	0.517	0.677
ABC	0.030	3	0.010	0.096	0.957
Error	5.633	54	0.104		
Analysis of Covariance					
Between Subjects					
A	0.004	1	0.004	0.024	0.851
Error	2.988	17	0.176		

TABLE 20

Analyses of Variance and Covariance on Mean Integrated Error on Both the X and Y Axes in the Time Sharing Task

Source of Variation	Sum of Squares	df	Mean Square	F Ratio	p Value
Analysis of Variance					
Between Subjects					
A (Noise)	39.379	1	39.379	0.302	0.595
Error	2346.287	18	130.349		
Within Subjects					
B (Session)	95.719	1	95.719	5.419	0.030
AB	8.377	1	8.377	0.474	0.506
Error	317.924	18	17.662		
C (Trial)	27.861	3	9.287	1.622	0.194
AC	1.746	3	0.582	0.102	0.954
Error	309.096	54	5.724		
BC	19.096	3	6.365	2.273	0.089
ABC	1.391	3	0.464	0.166	0.916
Error	151.201	54	2.800		
Analysis of Covariance					
Between Subjects					
A	31.687	1	31.687	0.687	0.576
Error	784.526	17	46.149		

TABLE 21

Analyses of Variance and Covariance on Mean Integrated Error on Both the X and Y Axes on the Rate Control Task

Source of Variation	Sum of Squares	df	Mean Square	F Ratio	p Value
Analysis of Variance					
Between Subjects					
A (Noise)	1.920	1	1.920	0.050	0.809
Error	686.170	18	38.121		
Within Subjects					
B (Session)	44.672	1	44.672	3.859	0.062
AB	1.441	1	1.441	0.125	0.726
Error	208.359	18	11.576		
C (Trial)	28.916	3	9.639	2.690	0.054
AC	10.662	3	3.554	0.992	0.595
Error	193.471	54	3.583		
BC	5.795	3	1.932	1.147	0.339
ABC	6.334	3	2.111	1.254	0.299
Error	90.951	54	1.684		
Analysis of Covariance					
Between Subjects					
A	3.349	1	3.349	0.140	0.712
Error	405.734	17	23.867		

TABLE 22

Analyses of Variance and Covariance on Mean Integrated Error on the X-Axis in the Time Sharing Task

Source of Variation	Sum of Squares	*df*	Mean Square	*F* Ratio	*p* Value
		Analysis of Variance			
Between Subjects					
A (Noise)	82.135	1.	82.135	0.600	0.546
Error	2463.283	18.	136.849		
Within Subjects					
B (Session)	94.615	1.	94.615	6.146	0.022
AB	7.992	1.	7.992	0.519	0.513
Error	277.088	18.	15.394		
C (Trial)	24.787	3.	8.262	1.365	0.262
AC	2.854	3.	0.951	0.157	0.922
Error	326.758	54.	6.051		
BC	24.258	3.	8.086	2.511	0.067
ABC	0.986	3.	0.329	0.102	0.954
Error	173.904	54.	3.220		
		Analysis of Covariance			
Between Subjects					
A	54.321	1.	54.321	1.070	0.316
Error	863.301	17.	50.782		

TABLE 23

Analyses of Variance and Covariance on Mean Integrated Error on the X-Axis in the Rate Control Task

Source of Variation	Sum of Squares	df	Mean Square	F Ratio	p Value
	Analysis of Variance				
Between Subjects					
A (Noise)	0.896	1.	0.896	0.021	0.857
Error	759.526	18.	42.196		
Within Subjects					
B (Session)	51.133	1.	51.133	4.189	0.053
AB	0.000	1.	0.000	0.000	1.000
Error	219.709	18.	12.206		
C (Trial)	30.166	3.	10.055	2.726	0.052
AC	12.178	3.	4.059	1.101	0.357
Error	199.176	54.	3.688		
BC	8.586	3.	2.862	1.384	0.257
ABC	8.443	3.	2.814	1.361	0.264
Error	111.658	54.	2.068		
	Analysis of Covariance				
Between Subjects					
A	0.443	1.	0.443	0.020	0.860
Error	383.211	17.	22.542		

TABLE 24

Analyses of Variance and Covariance on Mean Integrated Error on the Y-Axis in the Time Sharing Task

Source of Variation	Sum of Squares	df	Mean Square	F Ratio	p Value
			Analysis of Variance		
Between Subjects					
A (Noise)	12.191	1.	12.191	0.093	0.758
Error	2372.348	18.	131.797		
Within Subjects					
B (Session)	96.824	1.	96.824	4.679	0.042
AB	8.781	1.	8.781	0.424	0.529
Error	372.481	18.	20.693		
C (Trial)	32.590	3.	10.863	1.669	0.183
AC	1.047	3.	0.349	0.054	0.978
Error	351.410	54.	6.508		
BC	17.484	3.	5.828	1.620	0.194
ABC	3.582	3.	1.194	0.332	0.804
Error	194.234	54.	3.597		
			Analysis of Covariance		
Between Subjects					
A	14.728	1.	14.728	0.313	0.590
Error	801.063	17.	47.121		

TABLE 25

Analyses of Variance and Covariance on Mean Integrated Error on the Y-Axis in the Rate Control Task

Source of Variation	Sum of Squares	df	Mean Square	F Ratio	p Value
		Analysis of Variance			
Between Subjects					
A (Noise)	3.367	1.	3.367	0.082	0.769
Error	741.240	18.	41.180		
Within Subjects					
B (Session)	38.672	1.	38.672	2.510	0.127
AB	5.516	1.	5.516	0.358	0.563
Error	277.287	18.	15.405		
C (Trial)	28.332	3.	9.444	2.355	0.081
AC	11.320	3.	3.773	0.941	0.571
Error	216.545	43.	4.010		
BC	4.801	3.	1.600	0.831	0.515
ABC	7.563	3.	2.521	1.309	0.280
Error	104.006	54.	1.926		
		Analysis of Covariance			
Between Subjects					
A	8.515	1.	8.515	0.285	0.606
Error	507.904	17.	29.877		

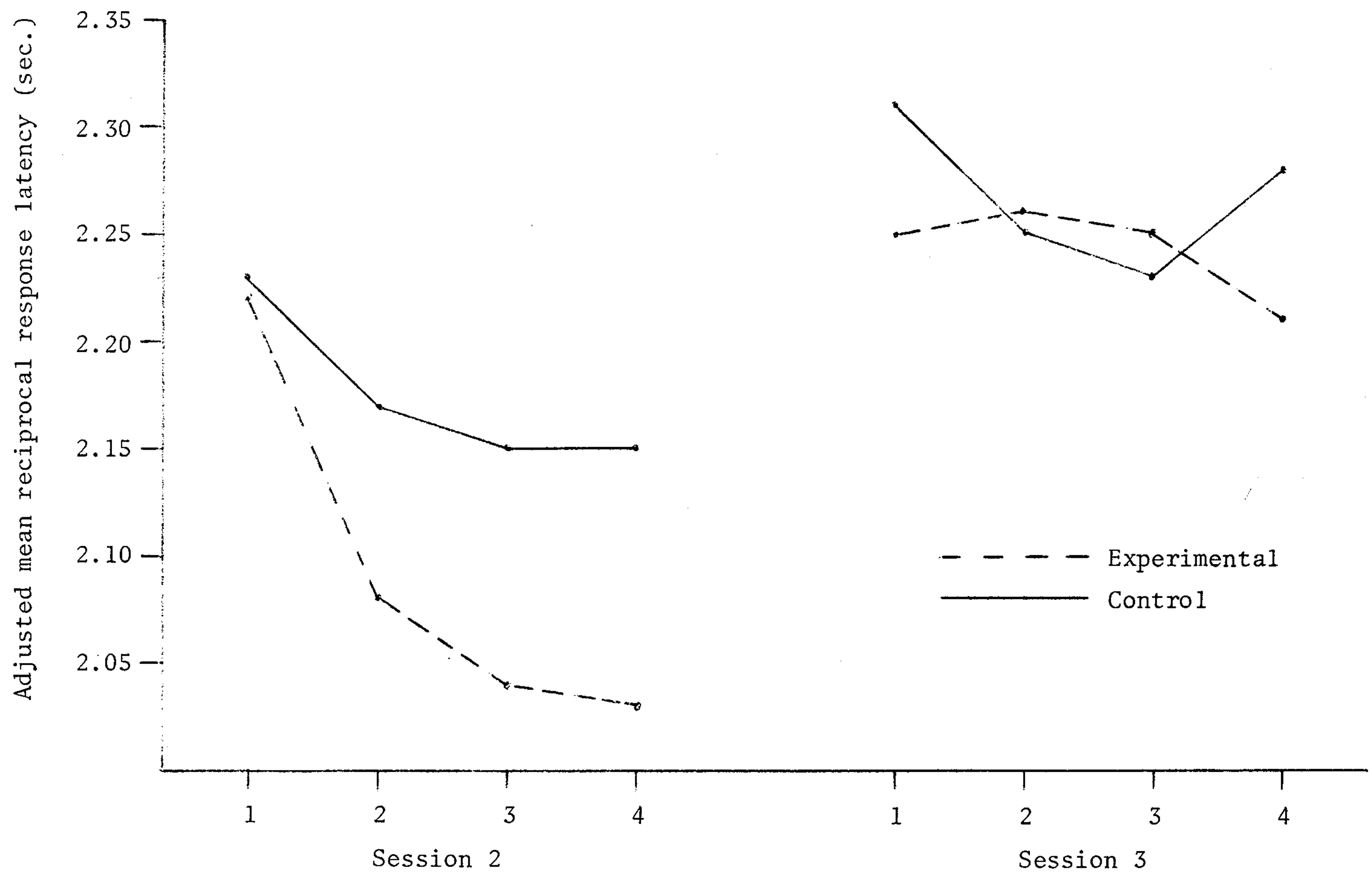


Fig. 3. Adjusted transformed mean response latency at each block of trials in both sessions on the Time Sharing task.

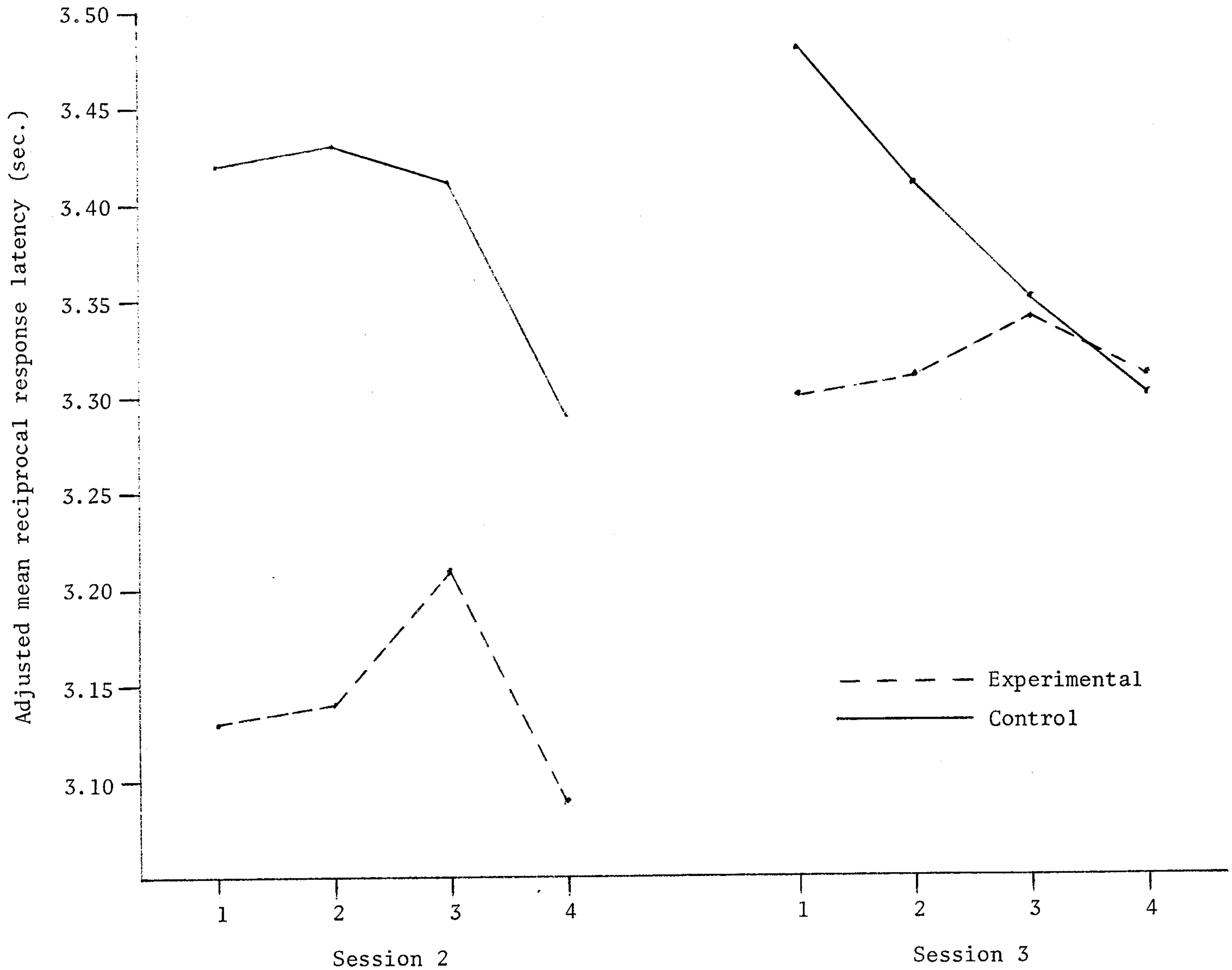


Fig. 4. Adjusted transformed mean response latency at each block of trials in both sessions on the Reaction Time task.

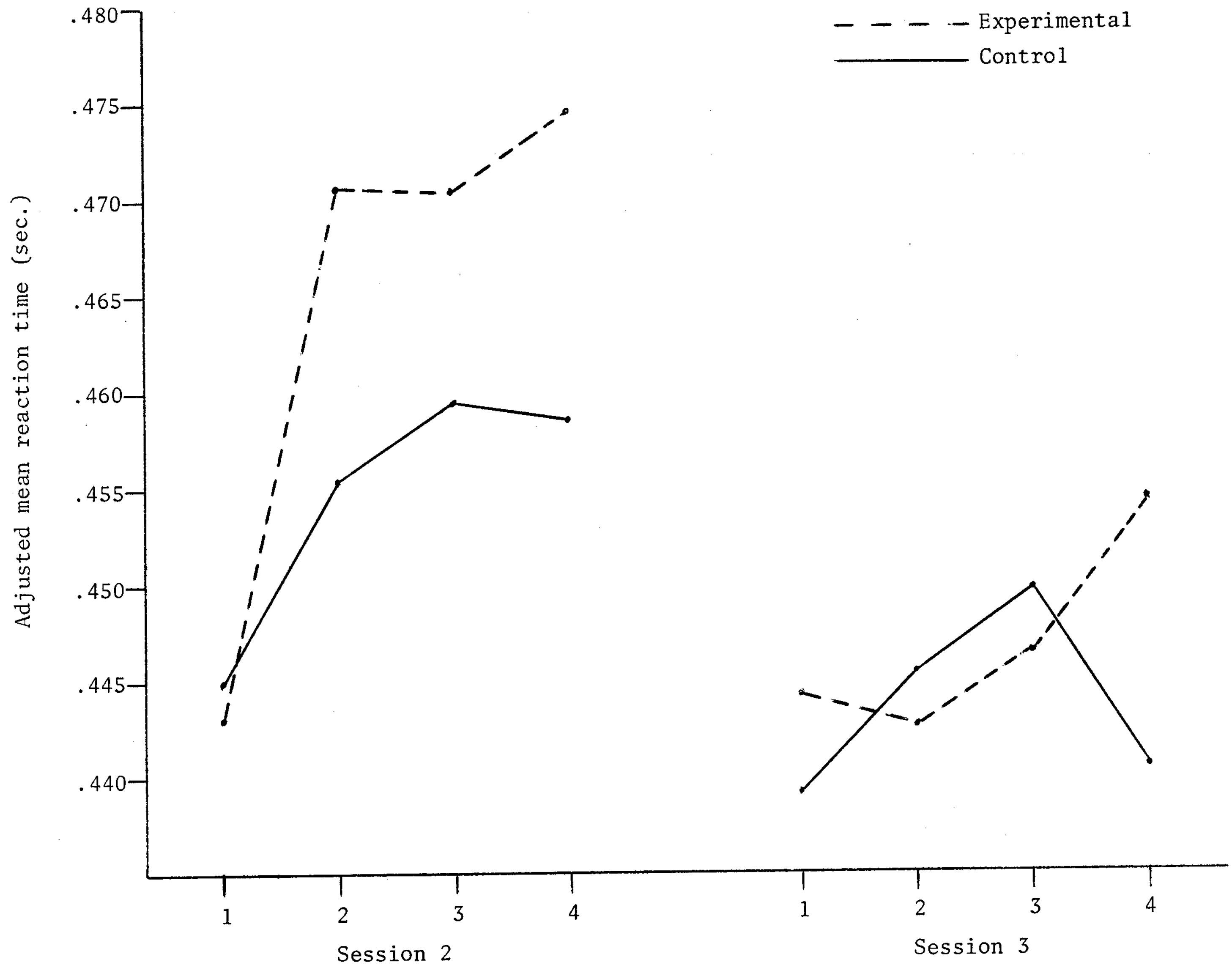


Fig. 5. Adjusted mean reaction time at each block of trials in both sessions on the Time Sharing task.

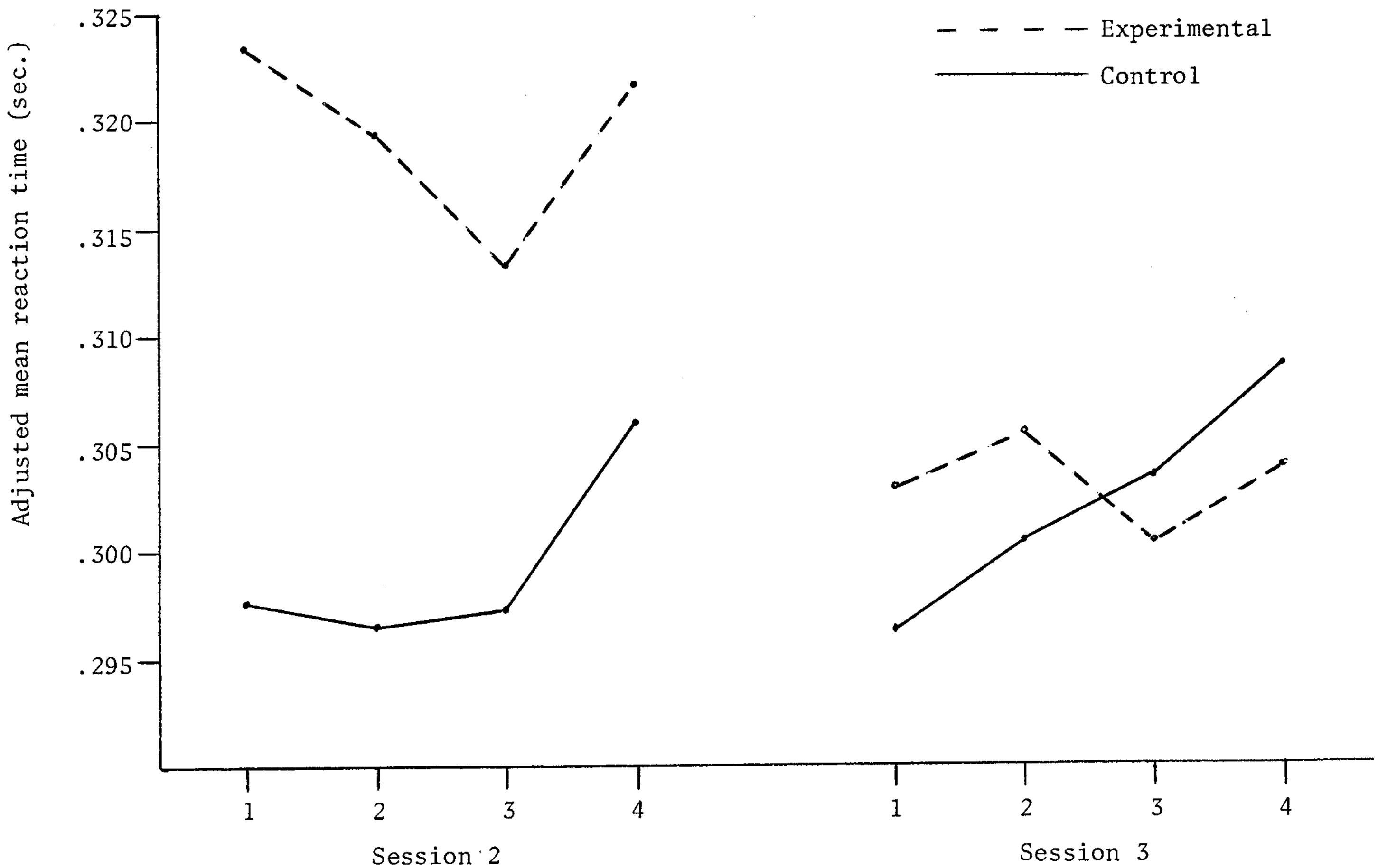


Fig. 6. Adjusted mean reaction time at each block of trials in both sessions on the Reaction Time task.

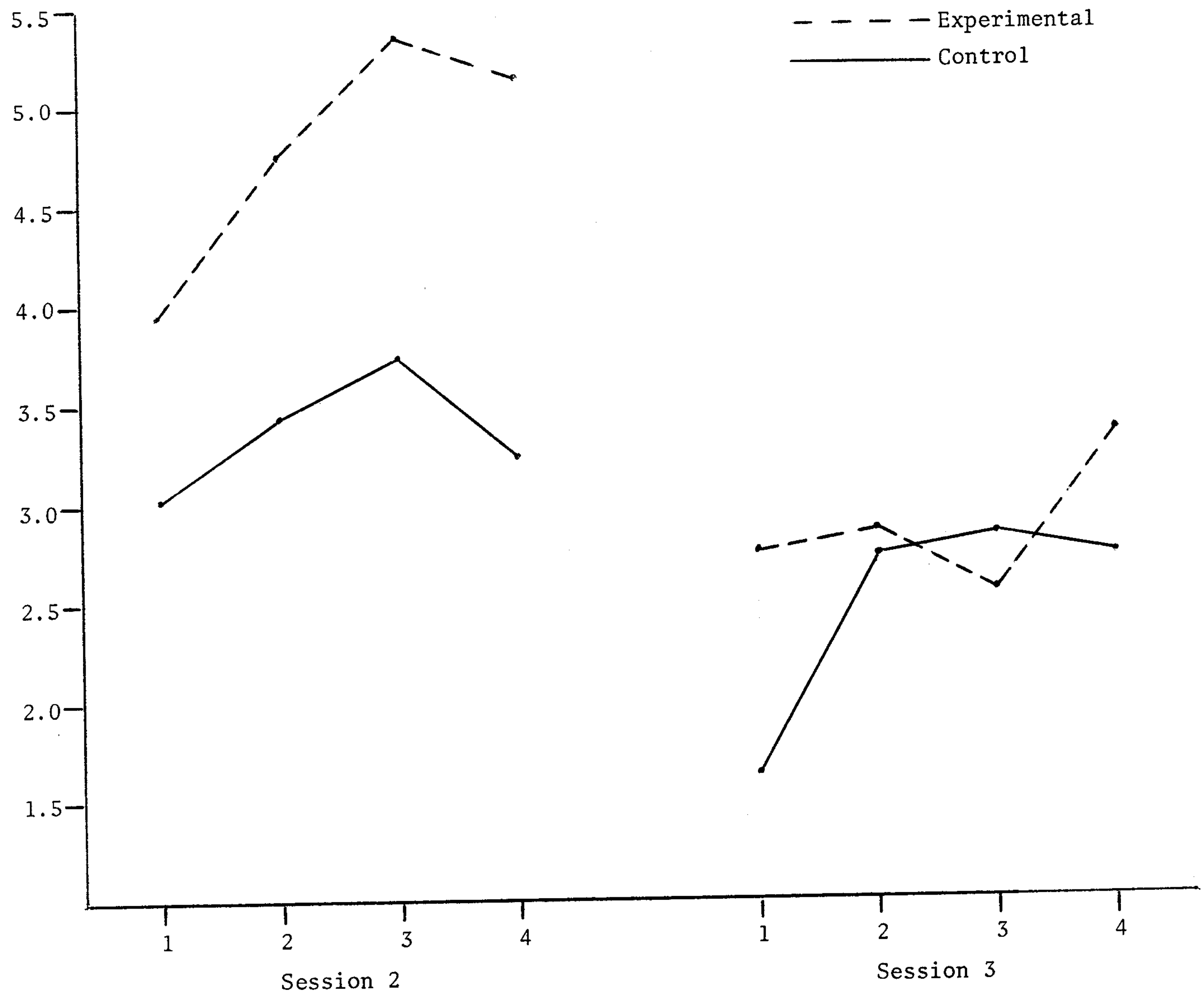


Fig. 7. Adjusted mean frequency of response blocks at each block of trials in both sessions on the Time Sharing task.

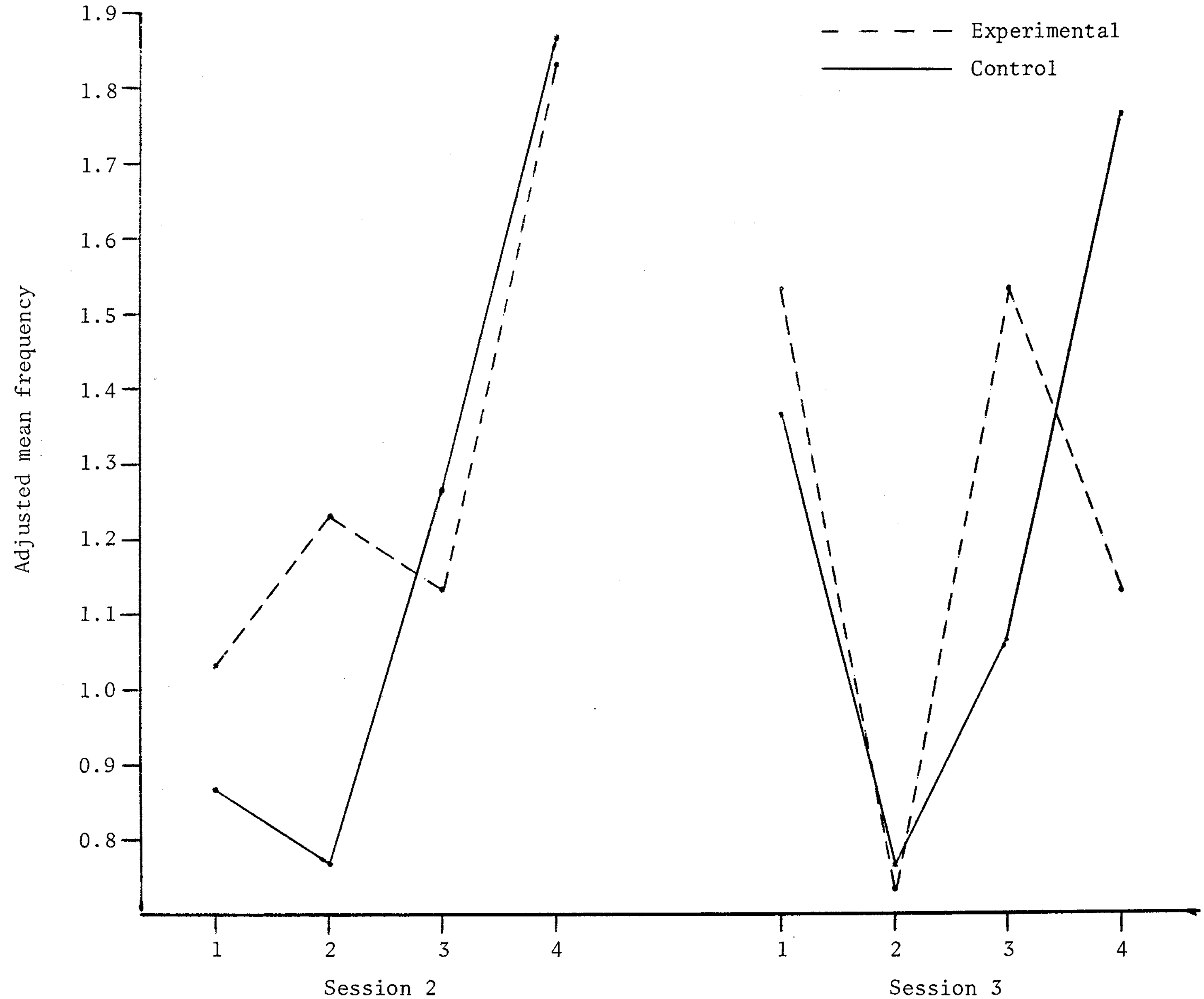


Fig. 8. Adjusted mean frequency of response blocks at each block of trials in both sessions on the Reaction Time task.

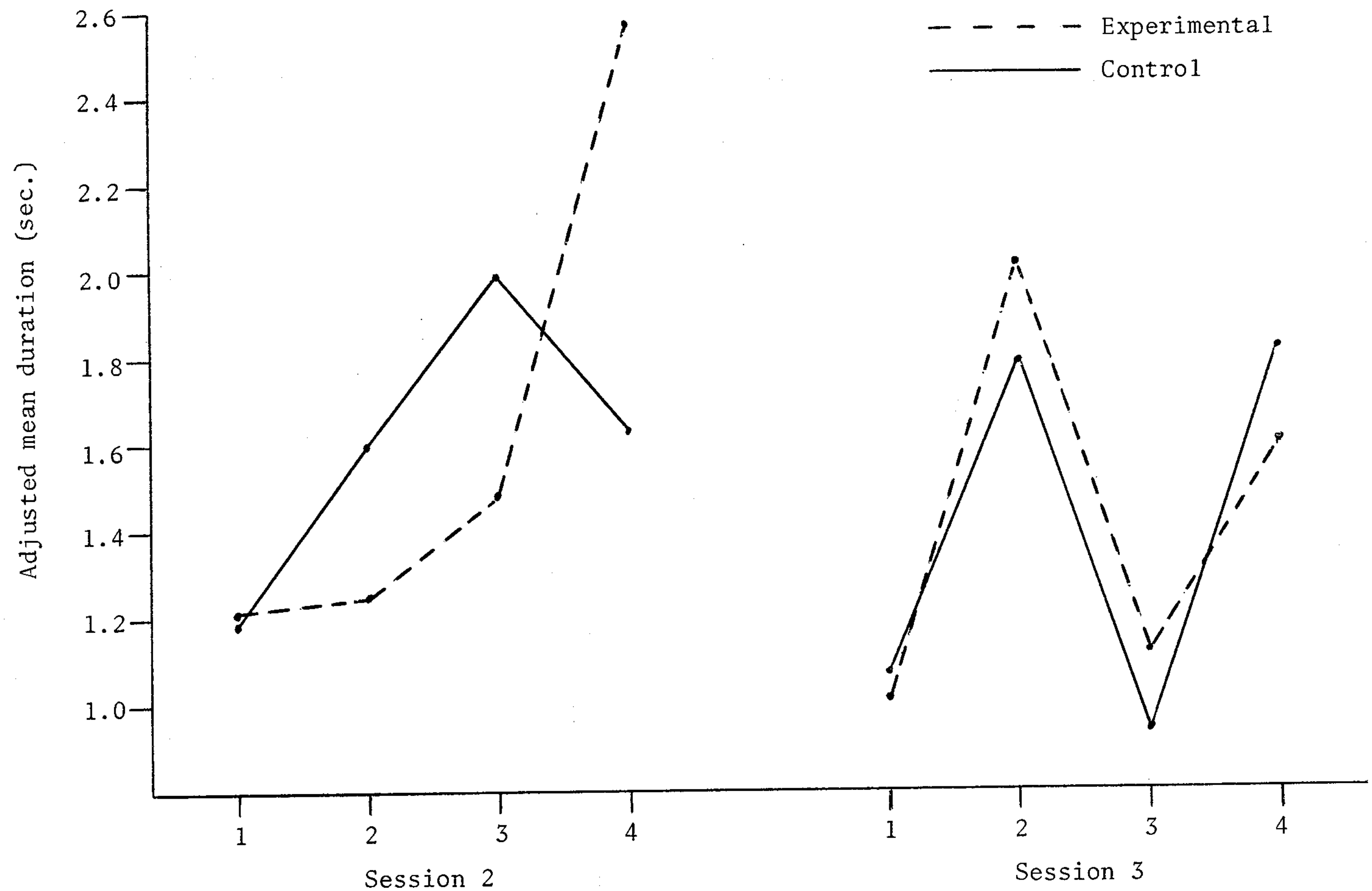


Fig. 9. Adjusted mean duration of a response block at each block of trials in both sessions on the Time Sharing task.

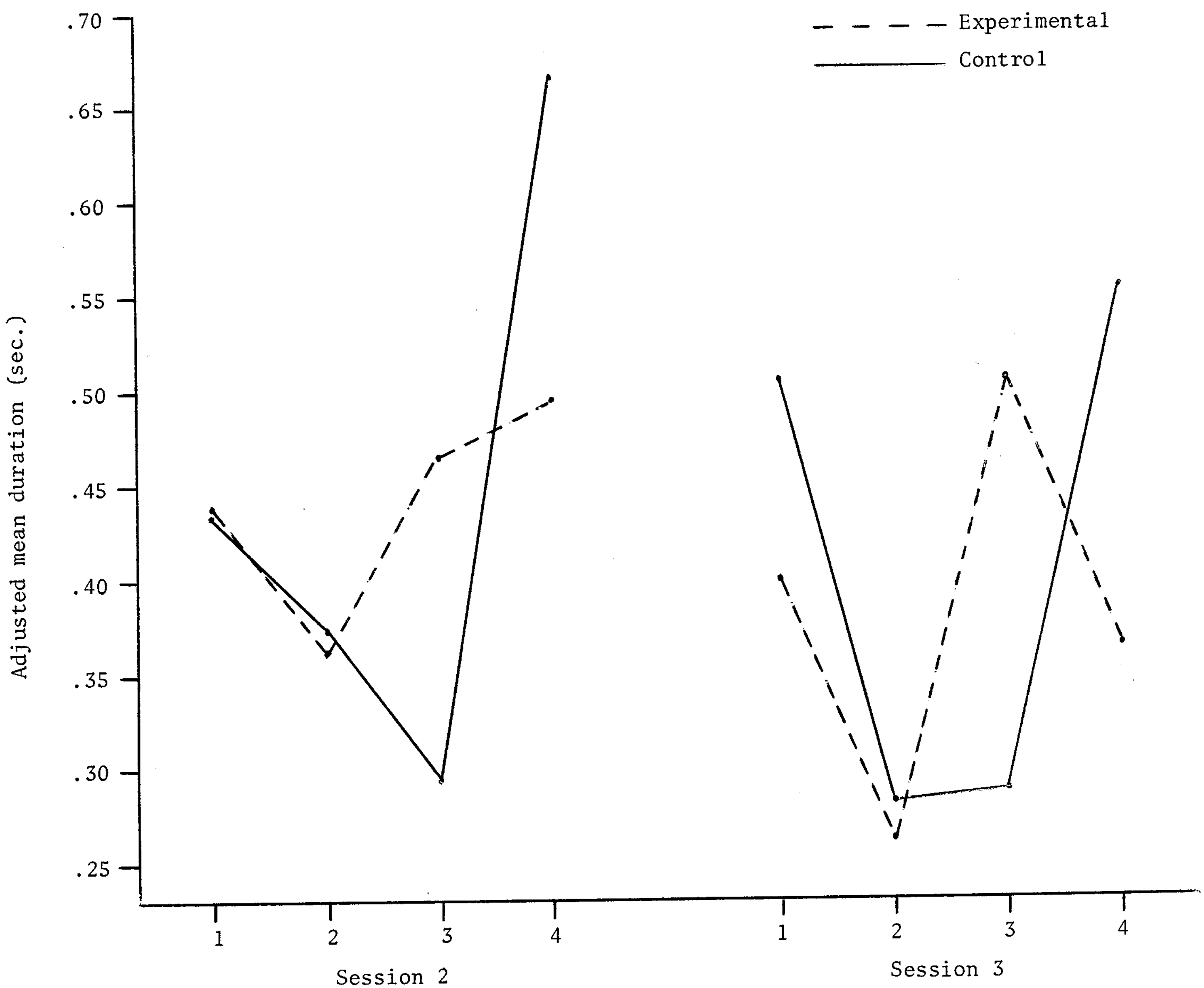


Fig. 10. Adjusted mean duration of a response block at each block of trials in both sessions on the Reaction Time task.

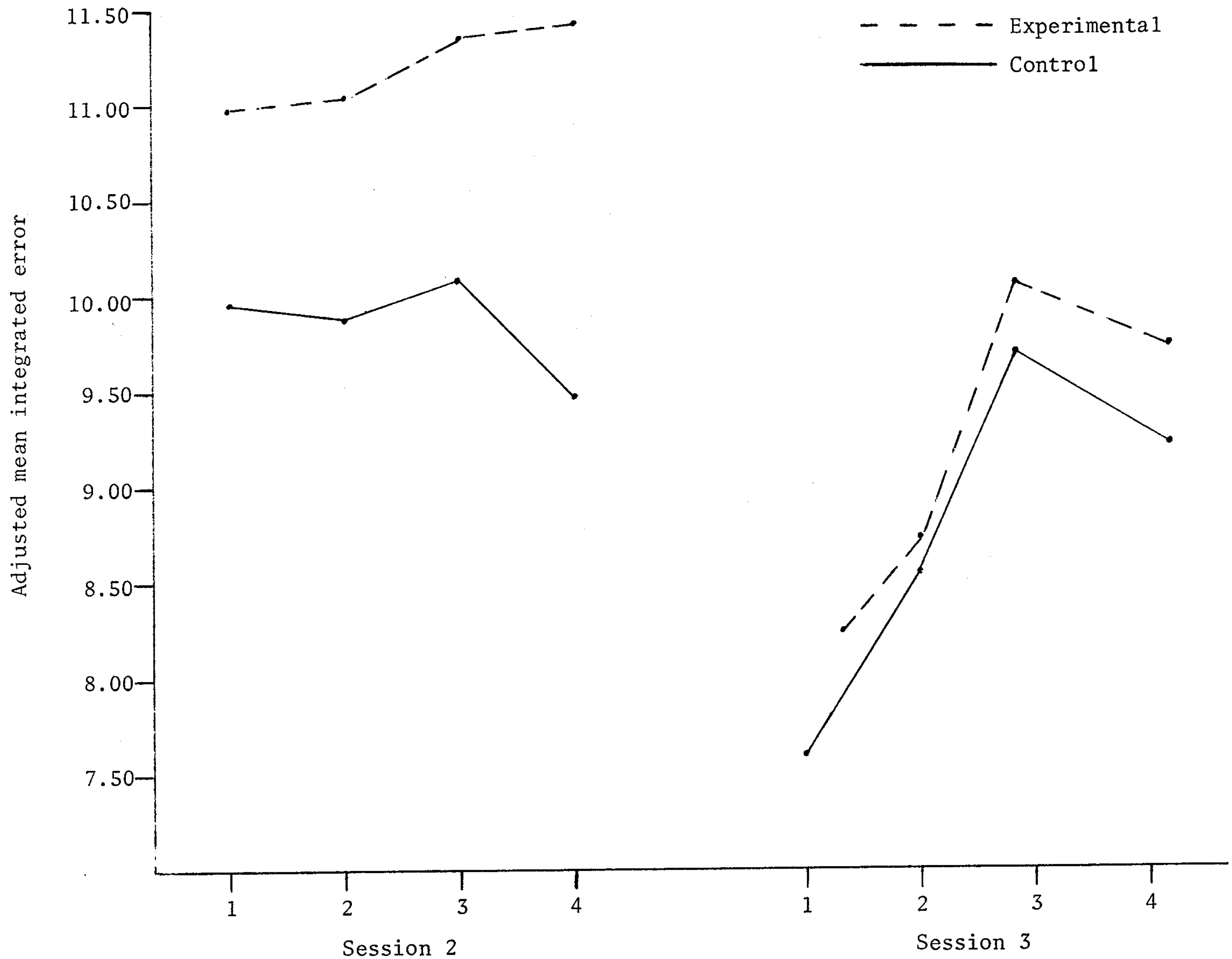


Fig. 11. Adjusted mean integrated error on both the x and y axes at each block of trials in both sessions on the Time Sharing task.

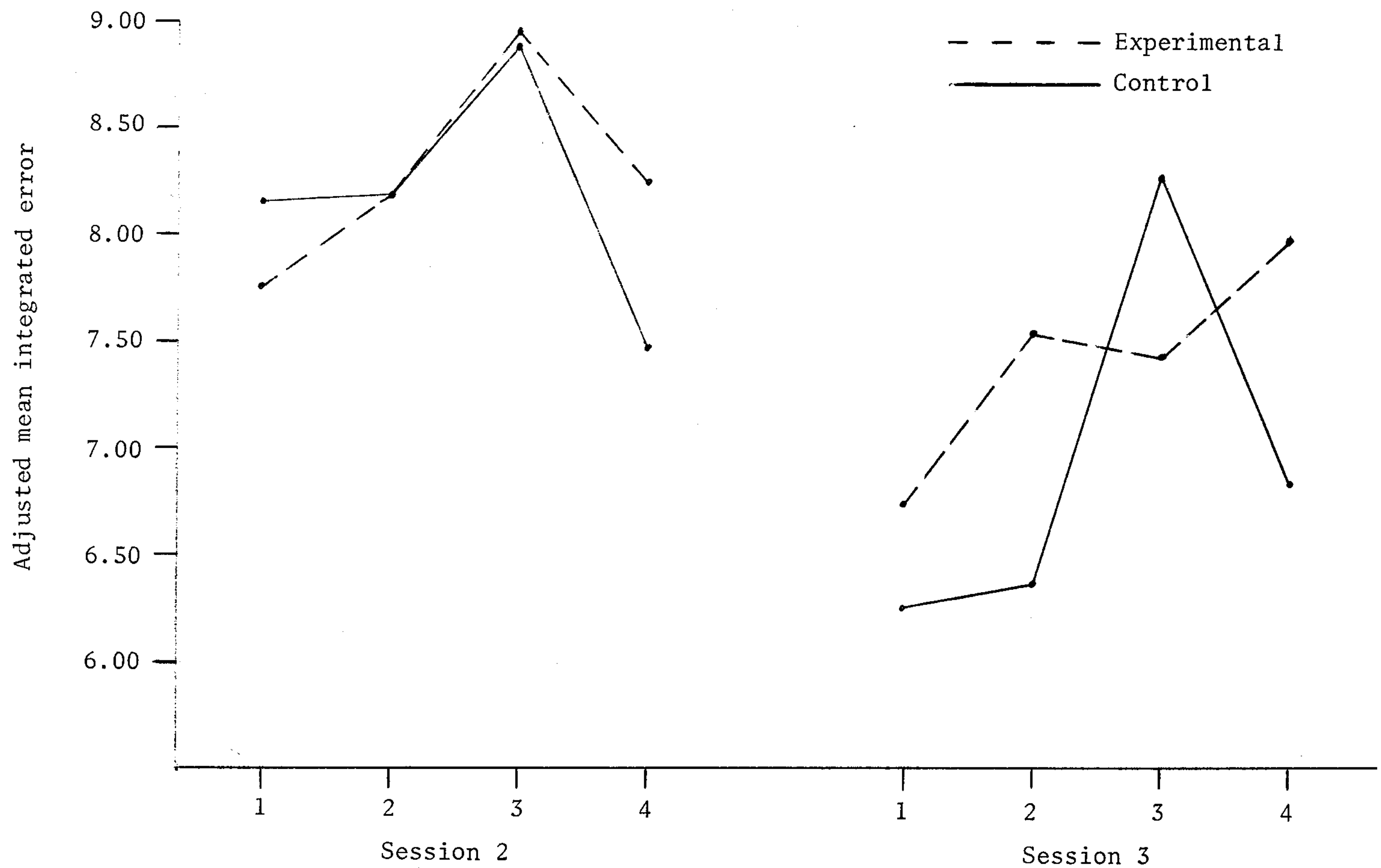


Fig. 12. Adjusted mean integrated error on both the x and y axes at each block of trials in both sessions on the Rate Control task.

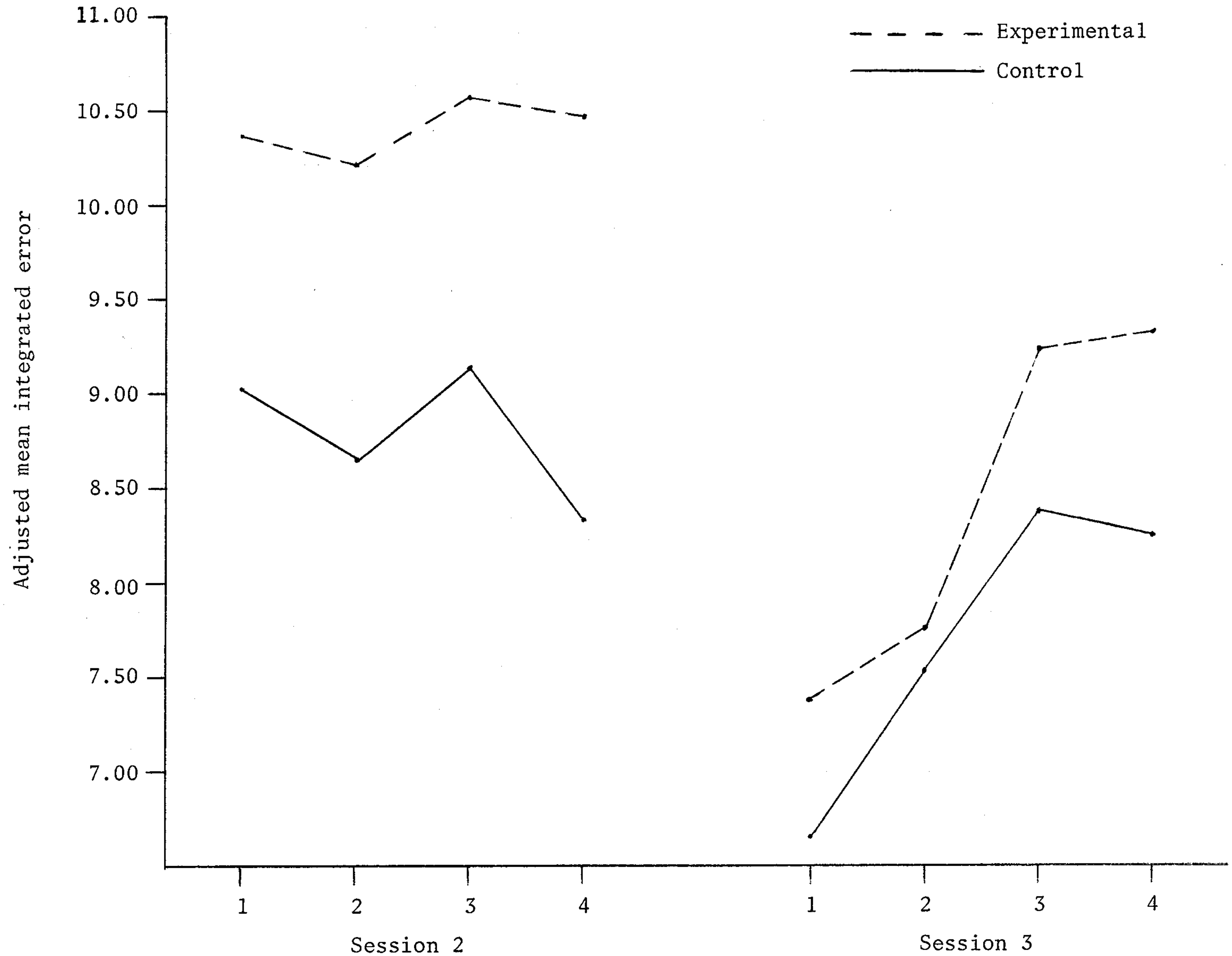


Fig. 13. Adjusted mean integrated error on the x-axis at each block of trials in both sessions on the Time Sharing task.

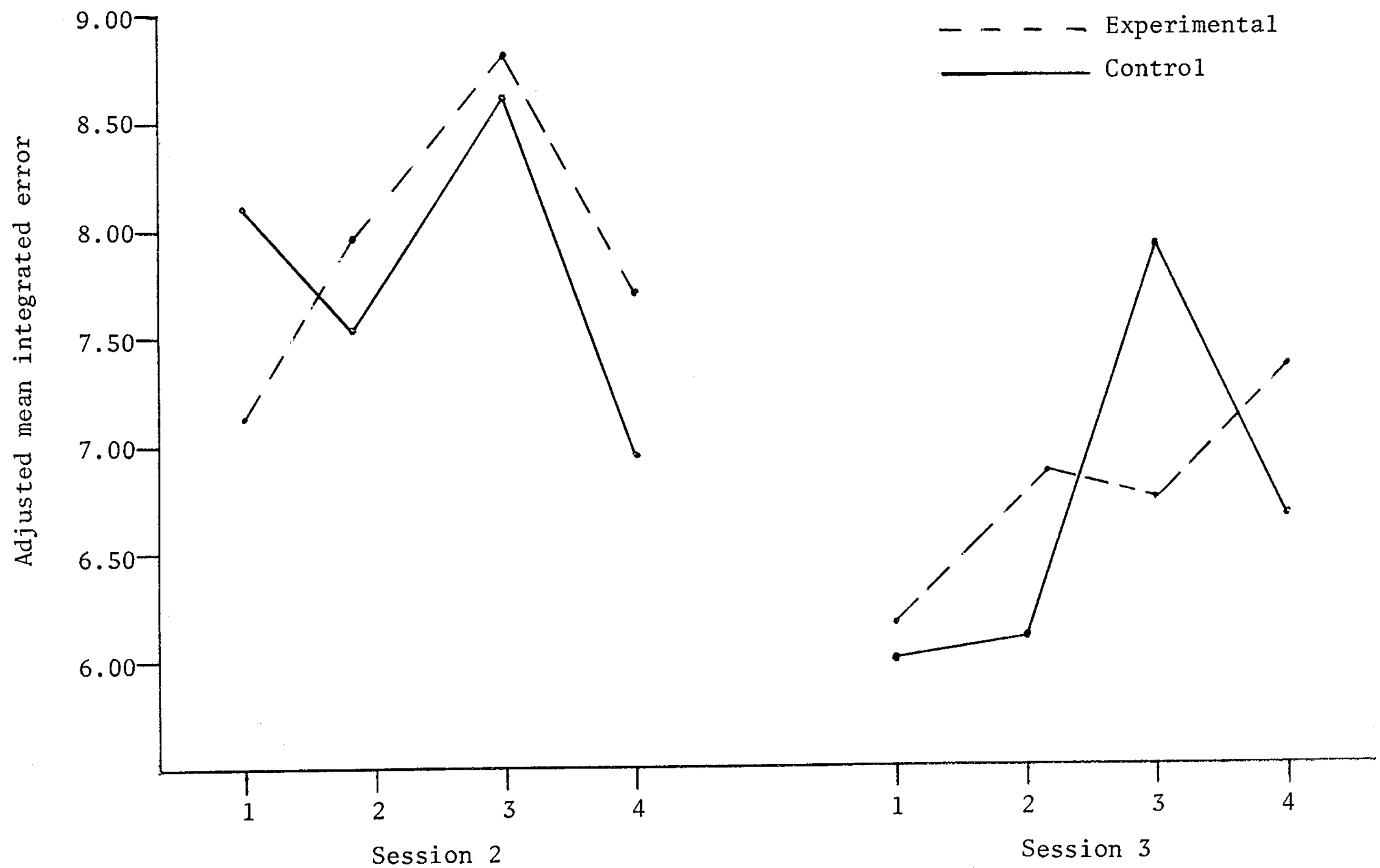


Fig. 14. Adjusted mean integrated error on the x-axis at each block of trials in both sessions on the Rate Control task.

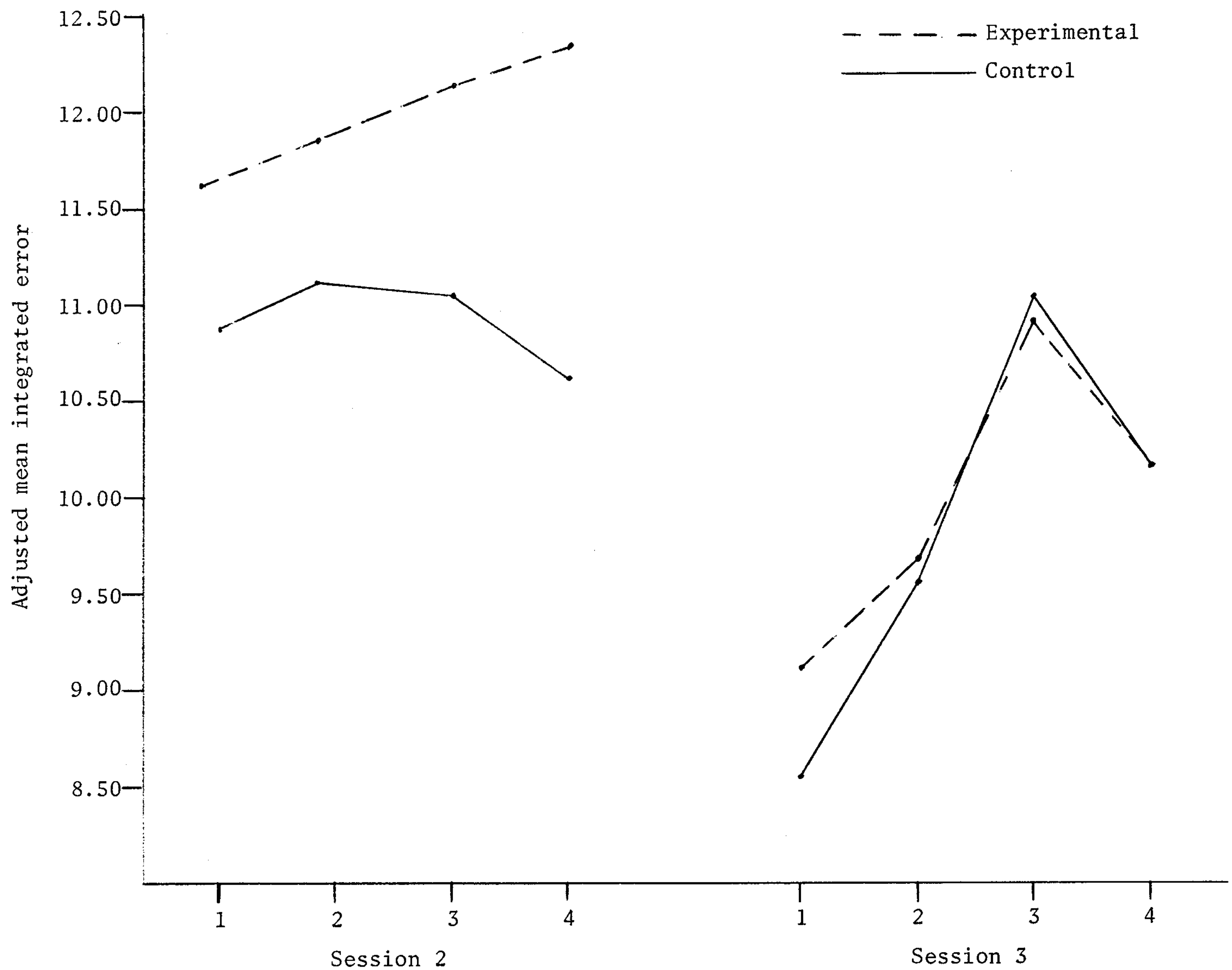


Fig. 15. Adjusted mean integrated error on the y-axis at each block of trials in both sessions on the Time Sharing task.

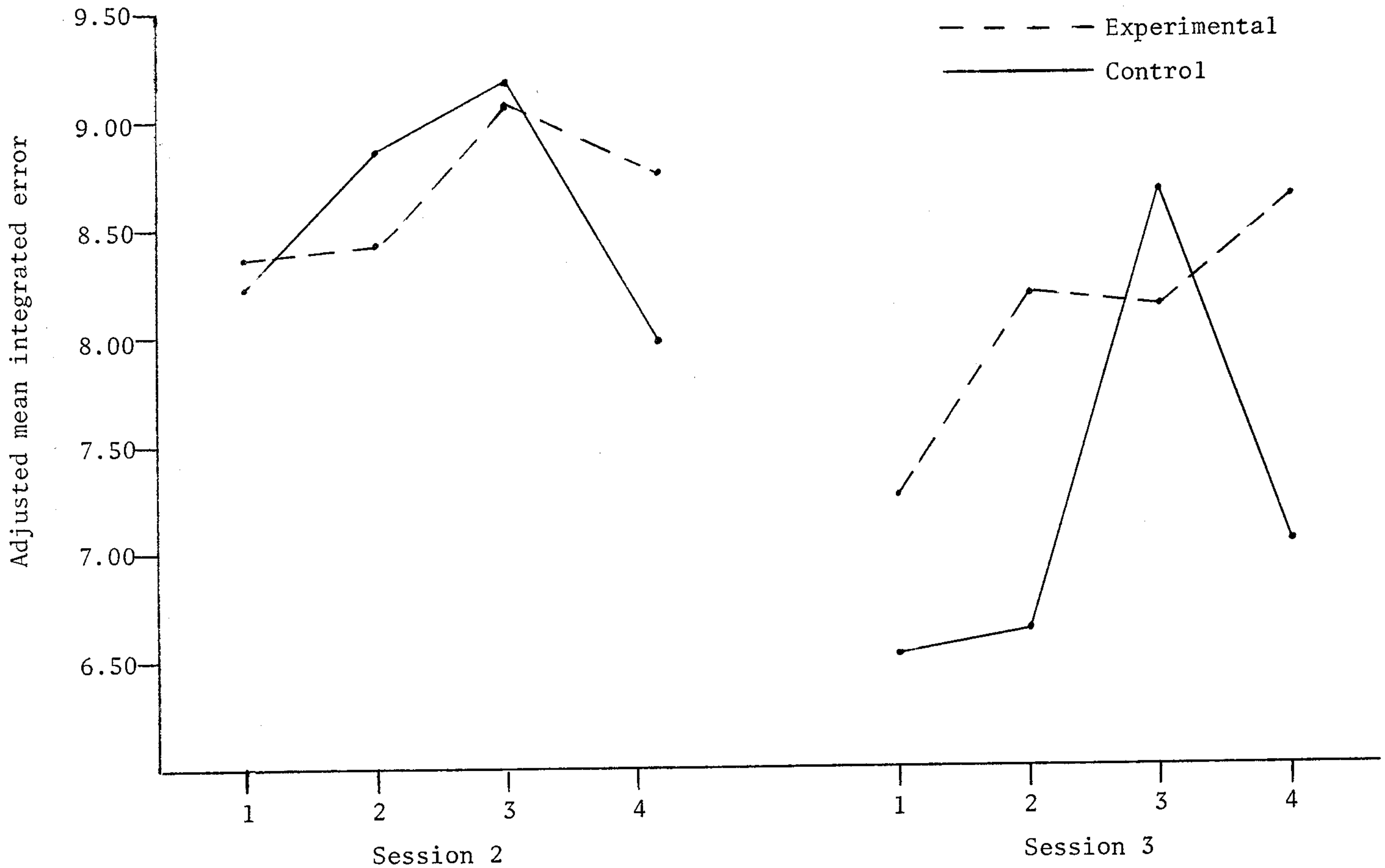


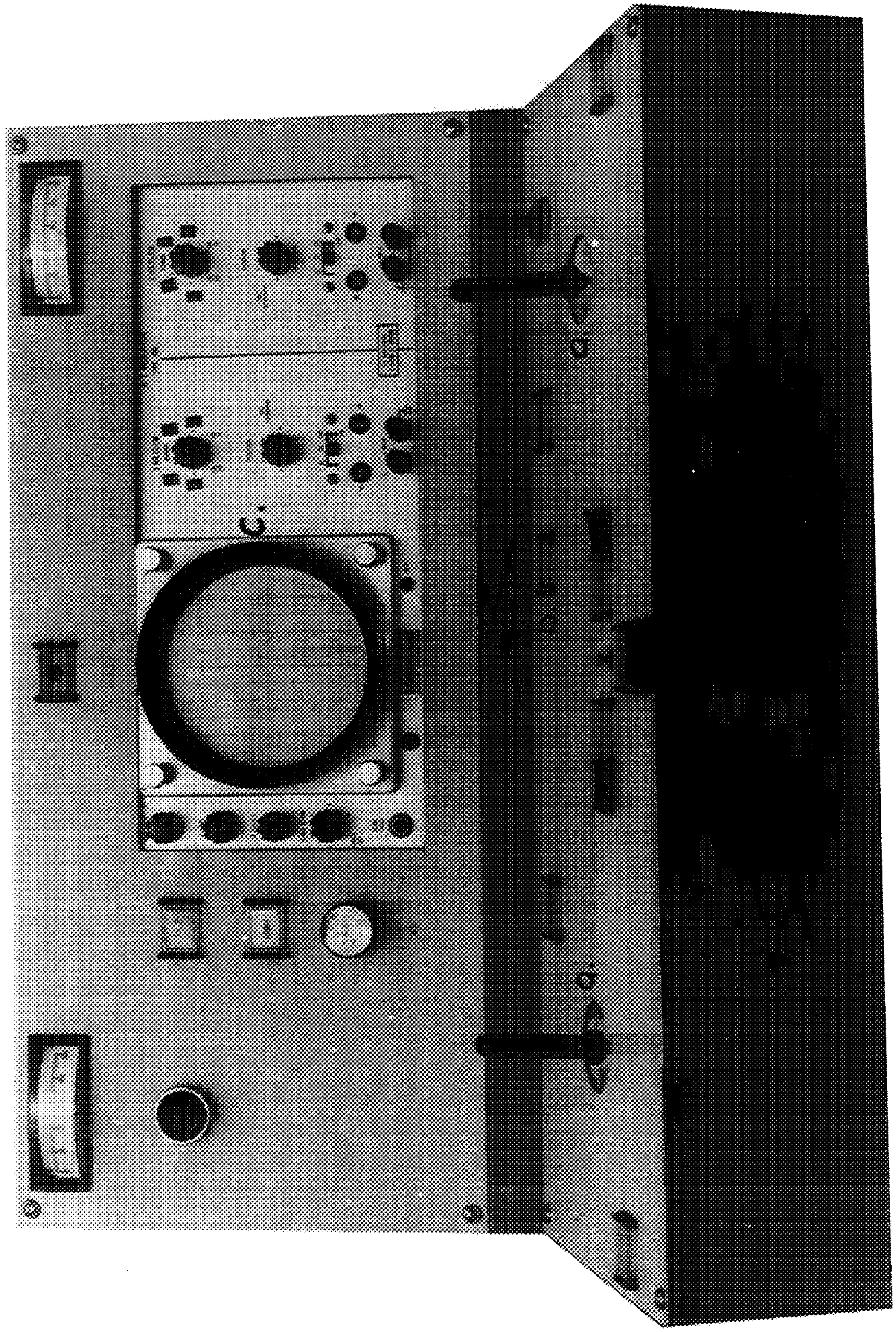
Fig. 16. Adjusted mean integrated error on the y-axis at each block of trials in both sessions on the Rate Control task.

APPENDIX D

Perceptual-Motor Console (PEMCON)--Operator's Console

(Originally described in Parker, Reilly, Dillon, Andrews, & Fleishman, 1965.)

The portions of the device employed in the present study included: a. the joysticks (left or right, depending upon subject's handedness); b. a backlighted reaction time key to which the subjects responded when the backlight came on; and c. the CRT screen which subjects observed for target movement while tracking with one of the joysticks.



CR-2149 – 4

FIRST CLASS MAIL

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION OFFICE

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Washington, D.C. 20546